

# **Social and Spatial Determinants of Adverse Birth Outcome Inequalities in Socially Advanced Societies**

by

Gang Meng

A thesis  
presented to the University of Waterloo  
in fulfillment of the  
thesis requirement for the degree of  
Doctor of Philosophy  
in  
Planning

Waterloo, Ontario, Canada, 2010

© Gang Meng 2010

## **AUTHOR'S DECLARATION**

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

## **Abstract**

The incidence of adverse birth outcomes, such as low birth weight and preterm births, has steadily risen in recent years in Canada. Despite the fact that numerous individual and neighbourhood risk factors for low birth weight and preterm births have been identified and various person-oriented intervention strategies have been implemented, uncertainties still exist concerning the role that place and space play in determining adverse birth outcomes.

In order to succeed in producing community-oriented health policy and planning guidelines to reduce both the occurrence and inequalities of adverse birth outcomes, the research presented in this thesis provides an approach to examining the pathways of various socio-economic, environmental, and psycho-social risks to LBW and preterm births. Using a modified multilevel binary-outcome mediational analysis method, case studies are conducted within three public health units in Ontario, namely the Wellington-Dufferin-Guelph Health Unit, the Windsor-Essex County Health Unit, and the Halton Region Health Unit. Different pathways are investigated given the available data and the theoretical assumptions of three health inequality pathway models, namely the behavioural model, the psycho-social model, and the materialist model, and the geographical and planning perspectives of health inequalities. A local spatial analysis process is also used to identify spatial clusters of incidence and to assess possible associated reasons in order to support public health policies and planning in community-oriented health interventions. Using Bayesian spatial hierarchical analysis and spatial clustering analysis, local clustering of high risks of adverse birth outcomes and spatial variations of associated individual risks within the study areas are identified.

The analysis is framed around five hypotheses that examine personal vs. spatial, compositional vs. contextual, psycho-social vs. material, personal vs. cultural, and global vs. local effects on the determinants of adverse birth outcomes. The results of testing these hypotheses provide evidence to assist with multi-component multi-level community-oriented interventions. Possible improvements of current prenatal care policies and programs to reduce the spatial and social inequalities of adverse birth outcomes are suggested. Potential improvements, including early stage prenatal health education, local healthy food provision, and cross-sector interventions such as the combination of social mixing strategies with bottom-up community-based health promotion programs, are also suggested.

## Acknowledgements

I would like to express my sincere gratitude to many people for their support and assistance. The completion of this thesis could not have been achieved without their support.

First of all, I would like to extend my sincere thanks to my research supervisors, Professors G. Brent Hall and Mary E. Thompson, for their unconditional time, advice and dedication in supervising my research work. Their expertise in various aspects of health geography and health statistics has been an important source of inspiration and motivation throughout the course of my study. I greatly appreciate their ways of supervising and questioning. They always put me back on the right track so that I can continually move forward confidently whilst allowing me the room to think independently. It is my great luck to be their student and I will cherish in my heart the relation with them.

I am thankful to Dr. Patrick Seleski for his academic advices and sustaining supports both in terms of research collaboration building and data collection. I am also thankful to Dr. Jane Law for her great advice in Bayesian statistical modeling and generous research funding support in the later years of my thesis work. Thanks are also extended to Dr. Ian McKillop for taking his time to read the thesis and provide valuable feedbacks through the course of my thesis work.

Many thanks are also due to the Wellington-Dufferin-Guelph Public Health Unit, Halton Region Public Health Unit and Windsor and Essex Country Public Health Unit for access to the data on births and to The South-Western Ontario Research Data Centre of Statistics Canada for access to the CCHS and Census data.

Finally, I am deeply thankful to my wife, Jilan Yang, for her enduring love, encouragement and support during my life, and my son, Toby Meng, for the great joy that he gives me after he came to the world in the third year of my PhD study.

## Table of Contents

AUTHOR’S DECLARATION.....	ii
Abstract.....	iii
Acknowledgements .....	iv
Table of Contents.....	v
List of Figures.....	ix
List of Tables .....	xii
Chapter 1 Introduction.....	1
1.1 Research gaps .....	4
1.2 Research goals and objectives .....	5
1.3 Organization of the thesis .....	7
Chapter 2 Determinants of Adverse Birth Outcomes .....	9
2.1 Definitions and organizing framework.....	9
2.2 Intrapersonal risks.....	11
2.2.1 Maternal age .....	12
2.2.2 Maternal Anthropometric factors.....	13
2.2.3 Genetic factors.....	14
2.2.4 Maternal health and medical conditions .....	14
2.2.5 Multiple births .....	15
2.2.6 Maternal under-nutrition.....	15
2.2.7 Maternal infections .....	16
2.2.8 Maternal stress and depression .....	17
2.2.9 Life style and health behaviours .....	18
2.3 Interpersonal risks.....	21
2.3.1 Marital status .....	21
2.3.2 Family violence or abuse .....	21
2.4 Organizational risks .....	22
2.4.1 Occupational Factors .....	22
2.5 Community and society level risks.....	23
2.5.1 Environmental pollution .....	24
2.5.2 Race/ethnicity (culture vs. personality) .....	25
2.5.3 Socio-economic status .....	26

2.5.4 Neighbourhood conditions.....	27
2.5.5 Prenatal health care provision and education .....	28
2.5.6 Social capital.....	31
2.6 Summary.....	32
Chapter 3 Theoretical Approaches to Studying Inequalities of Health and Adverse Birth Outcomes .....	34
3.1 Theoretical models of health inequalities .....	34
3.1.1 The behavioural model .....	35
3.1.2 The psycho-social model .....	36
3.1.3 The materialist and neo-materialist model.....	37
3.1.4 The life-course approach .....	39
3.2 Debates on composition and context .....	44
3.3 Geographical and planning perspectives on health inequalities .....	46
3.3.1 Ecological landscapes.....	46
3.3.2 Theories about space and place in social relationships.....	50
3.3.2.i Theories of structuration and the geography of power .....	51
3.3.2.ii Theories of political economy and the landscape of collective consumption .....	52
3.3.3 Social capital and community.....	56
3.4 Research hypotheses.....	58
3.5 Summary.....	63
Chapter 4 Methodology .....	64
4.1 Statistical analysis framework .....	64
4.2 Global analysis .....	66
4.2.1 Multilevel modeling.....	68
4.2.2 Mediation analysis .....	73
4.2.3 Testing of neighbourhood variations .....	77
4.2.3.i Intraclass correlation coefficient.....	77
4.2.3.ii Global spatial clustering .....	81
4.3 Local spatial analysis.....	83
4.3.1 Local clustering analysis.....	84
4.3.1.i Local indicators of spatial association .....	84
4.3.1.ii Spatially structured random effect models .....	85
4.3.1.iii Spatial scan statistics .....	91
4.3.2 Spatial regression analysis .....	93
4.3.2.i The geographically weighted regression method .....	94

4.3.2.ii Multilevel spatial models .....	95
4.3.2.iii Empirical Bayesian estimation for local variance .....	97
4.4 Preliminary analyses .....	98
4.4.1 Geocoding.....	99
4.4.1.i Geocoding by street address .....	99
4.4.1.ii Geocoding by postal code.....	100
4.4.2 Spatial interpolation.....	101
4.4.3 Factor analysis .....	107
4.5 Summary.....	109
Chapter 5 Case Study .....	111
5.1 Description of the case study setting .....	111
5.1.1 Geographic and demographic context of the study region .....	111
5.1.2 Current health practices of the study region .....	113
5.2 Data Sources .....	115
5.2.1 Personal-level data extraction and management procedures .....	116
5.2.1.i Geocoding.....	117
5.2.2 Contextual-level data collection .....	118
5.2.2.i Spatial interpolation.....	119
5.3 Adverse birth outcome Analysis.....	126
5.3.1 Current adverse birth outcomes of the study region .....	126
5.3.2 Hypothesis 1: Personal vs. Spatial.....	128
5.3.3 Hypothesis 2: Compositional vs. Contextual Factors .....	133
5.3.3.i Personal and community-level risk identification .....	133
5.3.3.ii Factor analysis .....	139
5.3.3.iii Contextual effect testing.....	142
5.3.4 Mediation analysis on pathways of neighbourhood determinants of adverse birth outcomes .....	145
5.3.4.i Hypothesis 3: Psycho-social vs. Material Influences .....	145
5.3.4.ii Hypothesis 4: Personal vs. Cultural Influences .....	149
5.3.5 Hypothesis 5: Global vs. Local Influences .....	154
5.3.5.i Local spatial clustering analysis .....	155
5.3.5.ii Disease mapping.....	157
5.3.5.iii Spatial variation of personal risk factors .....	169
5.4 Discussion and recommendations on community-based intervention and health planning .....	179

5.5 Summary.....	185
Chapter 6 Conclusion .....	186
6.1 Thesis achievements .....	186
6.1.1 Theoretical achievements .....	186
6.1.2 Methodological improvements .....	188
6.1.3 Practical implications.....	189
6.2 Limitations and future research .....	191
6.2.1 Data constraints .....	191
6.2.2 Technical constraints .....	192
Appendix A Description of variable names.....	195
Appendix B SAS models for mediational analysis of LBW and preterm births .....	197
Appendix C Mediational analysis results for preterm births .....	199
Appendix D Mediational analysis results for LBW.....	203
Appendix E WinBUGS Code for the random effect models.....	208
Appendix F Model convergence testing for the random effect preterm birth model .....	210
Appendix G Spatial distribution of personal risk factors of adverse birth outcomes .....	213
References .....	233



## List of Figures

Figure 2.1: Social ecological model .....	10
Figure 3.1: Agent-host-environment Triad (Adapted from Young, 2005, p115) .....	34
Figure 3.2: The causal pathways of the behavioural explanation on health inequalities.....	36
Figure 3.3: The causal pathways of the psycho-social model (Adapted from Brunner and Marmot, 2006, p27).....	37
Figure 3.4: The causal pathways of the materialist/neo-materialist model .....	39
Figure 3.5: Life course models (where A, B, C, D denotes risk factors) (After Kuh and Ben-Shlomo, 2004, p10).....	40
Figure 3.6: An integrated framework of social determinants on health (After Marmot & Wilkinson, 2006, p9).....	44
Figure 3.7: The framework of human disease ecology (After Meade and Earickson, 2000, p25) .....	47
Figure 3.8: Black box epidemiology (After Curtis, 2004, p175).....	49
Figure 3.9: Conceptual framework for socially unequal environmental risk (Adapted from Curtis, 2004, p188).....	50
Figure 4.1: The quantitative procedure for adverse birth outcome analysis.....	65
Figure 4.2 A two-level structure of persons nested in neighbourhoods .....	68
Figure 4.3: Varying level-1 and level-2 relationships (After Subramanian et al., 2003, p71).....	71
Figure 4.4: A typical fitted semi-variogram model .....	105
Figure 5.1: Spatial locations and major cities of the three health units .....	112
Figure 5.2: The best fitted semi-variogram function for household income .....	121
Figure 5.3: Household income distribution interpolated using IDW with power 1.....	123
Figure 5.4: Household income distribution interpolated using Kriging .....	123
Figure 5.5: Aggregated CDA-level household income by IDW interpolation .....	124
Figure 5.6: Aggregated CDA-level household income by Kriging interpolation.....	124
Figure 5.7: Census 2006 CDA-level household income.....	125
Figure 5.8: Kriging standard errors for the household income variable .....	126
Figure 5.9: The spatial distribution of raw preterm birth rates .....	129
Figure 5.10: The spatial distribution of raw LBW rates .....	130
Figure 5.11: Preterm birth clusters in the study region.....	156
Figure 5.12: LBW birth clusters in the study region .....	156
Figure 5.13: Distribution of spatially correlated random effects $u_0$ for preterm births based on	

contiguity neighbourhood structure.....	162
Figure 5.14: Distribution of total random effects $u_0+v_0$ for preterm births based on a contiguity neighbourhood structure.....	163
Figure 5.15: Distribution of spatially correlated random effects $u_0$ for preterm births based on an inverse distance neighbourhood structure .....	164
Figure 5.16: Distribution of total effects $u_0+v_0$ for preterm births based on an inverse distance neighbourhood structure.....	165
Figure 5.17: Distribution of spatially correlated random effects $u_0$ for LBW births based on an contiguity neighbourhood structure.....	167
Figure 5.18: Distribution of total random effects $u_0+v_0$ for LBW births based on a contiguity neighbourhood structure.....	168
Figure 5.19: Traffic impacts on preterm birth clusters .....	174
Figure 5.20: Traffic impacts on LBW birth clusters.....	175
Figure 5.21: Total municipality level random effects of preterm births after controlling identified risks .....	177
Figure 5.22: Total municipality level random effects of LBW births after controlling identified risks .....	178
Figure G.1: Spatially correlated random effect $u_0$ for preterm birth.....	213
Figure G.2: Total random effects $u_0+v_0$ for preterm birth .....	214
Figure G.3: Spatially correlated random effect of Factor 1 (SES) $u_1$ for preterm birth.....	215
Figure G.4: Total random effects of Factor 1 (SES) $u_1+v_1$ for preterm birth.....	216
Figure G.5: Spatially correlated random effect of Factor 2 (psycho-social) $u_2$ for preterm birth.....	217
Figure G.6: Total random effects of Factor 2 (psycho-social) $u_2+v_2$ for preterm birth.....	218
Figure G.7: Spatially correlated random effect of Factor 3 (behavioural) $u_3$ for preterm birth.....	219
Figure G.8: Total random effects of Factor 3 (behavioural) $u_3+v_3$ for preterm birth .....	220
Figure G.9: Spatially correlated random effect of Factor 4 (health) $u_4$ for preterm birth.....	221
Figure G.10: Total random effects of Factor 4 (health) $u_4+v_4$ for preterm birth .....	222
Figure G.11: Spatially correlated random effect $u_0$ for LBW birth .....	223
Figure G.12: Total random effects $u_0+v_0$ for LBW birth.....	224
Figure G.13: Spatially correlated random effect of Factor 1 (behavioural) $u_1$ for LBW birth .....	225
Figure G.14: Total random effects of Factor 1 (behavioural) $u_1+v_1$ for LBW birth.....	226
Figure G.15: Spatially correlated random effect of Factor 2 (SES) $u_2$ for LBW birth .....	227
Figure G.16: Total random effects of Factor 2 (SES) $u_2+v_2$ for LBW birth.....	228
Figure G.17: Spatially correlated random effect of Factor 3 (psycho-social) $u_3$ for LBW birth .....	229

Figure G.18: Total random effects of Factor 3 (psycho-social) $u_3+v_3$ for LBW birth .....	230
Figure G.19: Spatially correlated random effect of Factor 4 (health) $u_4$ for LBW birth.....	231
Figure G.20: Total random effects of Factor 4 (health) $u_4+v_4$ for LBW birth.....	232

## List of Tables

Table 3.1: Comparison of influences, focuses, and policy implementation of health inequality models .....	41
Table 3.2: Advantages and disadvantages of health inequality models.....	42
Table 3.3: Compositional and contextual factors and influences .....	45
Table 3.4: Compositional and contextual explanations on health inequalities .....	45
Table 4.1: OLS and Multilevel binary outcome model comparison for mediational analysis .....	75
Table 5.1: Moran's I results for selected CCHS variables (Bold numbers represent statistically significant results at the 95% confidence level) .....	120
Table 5.2: Cross-validation comparison between models and parameters .....	121
Table 5.3: Preterm birth rate by birth year and health units (%) .....	127
Table 5.4: Low Birth Weight rate by birth year and health units (%) .....	127
Table 5.5: ICC estimates for LBW and preterm birth .....	128
Table 5.6: ICC estimates for LBW and preterm birth using random effect models .....	131
Table 5.7: Design effect.....	132
Table 5.8: Global Moran's I for the random effects of LBW and preterm births.....	132
Table 5.9: Regression results for preterm birth risks.....	135
Table 5.10: Regression results for LBW risks.....	137
Table 5.11: Factor pattern for preterm births.....	140
Table 5.12: Factor pattern for LBW .....	140
Table 5.13: The association of personal-level risk factors with preterm births.....	141
Table 5.14: The association of personal-level risk factors with LBW .....	142
Table 5.15: Contextual effects of neighbourhood variables on preterm births.....	143
Table 5.16: Contextual effects of neighbourhood variables on LBW .....	144
Table 5.17: Pearson Correlation Coefficients of neighbourhood ethnic variables and SES (variable names are defined in Appendix A).....	151
Table 5.18: Associations of ethnic composition and adverse birth outcomes after controlling for neighbourhood low SES .....	151
Table 5.19: Parameter estimates for preterm births based on a contiguity neighbourhood structure.....	160
Table 5.20: Parameter estimates for preterm births based on an IDW neighbourhood structure .....	160
Table 5.21: Parameter estimates for preterm births based on a contiguity neighbourhood structure.....	166
Table 5.22: Fixed and random effects of personal parameter estimates for preterm birth .....	170
Table 5.23: Fixed and random effect of personal parameter estimates for LBW birth .....	171

# **Chapter 1**

## **Introduction**

Over the centuries, despite overall improvements in health conditions and longevity, spatial and social disparities of the incidence and prevalence of chronic diseases have not lessened. Along with the globalization of the world's economy, urbanization and the polarization of socio-economic classes, marked social and geographical inequalities in health and disease have widened in almost all developed countries. These health gaps are especially remarkable between the socially advantaged and disadvantaged, or between the rich and the poor. This trend is unlikely to be overturned through market-driven economic growth. Health interventions to reduce these social and spatial health inequalities are therefore important not only for the sake of equity and reduction of the adverse impacts of social disparities, but also for the health improvement of all humans (Morone and Jacobs, 2005; Dreever et al., 1996).

Population health is determined by a number of hierarchical influences, from proximal personal-level biological, psychological and behavioural factors, to intermediate-level interpersonal and organizational factors (such as home and family influences, work and school settings, and civic organizations), to distal-level community and society determinants (Hanson et al., 2005). In the past, personal-level risk factors have been well recognized by health professionals and person-based health interventions were widely practiced. However, at the contextual-level, although social and environmental determinants are increasingly seen as the key determinants of health, and social and environmental inequalities of health have been observed at international, provincial, municipal, and even neighbourhood levels (Marmot and Wilkinson, 2006; Raphael, 2004), the exact role that these macro social and environmental determinants have played in shaping health inequality landscapes is still not clear. Few effective or targeted health interventions have been developed to address these inequalities. It is therefore necessary to study and reveal the mechanisms of social and environmental determinants of health and the role played by place and space within these mechanisms.

One good indicator of health inequalities is the inequalities of adverse birth outcomes. Adverse birth outcomes, such as low birth weight (LBW) and preterm births, not only are key outcomes of reproductive health and perinatal health, but also are closely linked to serious, both short- and long-term, physical and mental disabilities. Low birth weight babies due to intrauterine growth restriction are closely associated with perinatal and infant death, and a wide spectrum of physical and cognitive

disabilities and chronic health problems later in life, such as hypertension, ischemic heart disease, metabolic syndrome, stroke, diabetes, malignancies, osteoarthritis, and dementia (Barker et al., 1993; Barker et al., 1986). Preterm birth also is associated with increased risks in mortality and childhood morbidities such as developmental problems, cerebral palsy, learning difficulties, and hearing and visual impairments (McCormick, 1985; Kliegman et al., 1990). In addition, LBW and preterm babies also account for a disproportionately high percentage of the health care costs among all newborns, which put high economic burdens on hospital health care.

Just like most other health outcomes, adverse birth outcomes are consistently more frequent among socially and economically disadvantaged groups and societies (Urquia et al., 2007). Socio-economic disparities in adverse birth outcomes are one of the most persistent findings in perinatal research. Evidence of associations between low socio-economic status (mostly defined by one or a combination of low education, low income, and low occupation) and adverse birth outcomes exists across and within most countries (Kramer, et al., 2000), especially for groups with economic deprivation (Luo, et al., 2006; Krieger, et al., 2003). Because of its potential importance in the elimination of poverty, reduction of health inequalities, and improvements in people's lives, reduction in the prevalence of low birth weight is used as one important indicator for monitoring the reproductive health goals within the Millennium Development Goals (MDGs) identified at the Millennium Summit sponsored by the United Nations in September 2000.

Many programs to prevent adverse birth outcomes have been initiated and implemented in developed countries. These include medical screening and treatment, prenatal care, prenatal education, and various types of financial assistance and social support. However, although dramatic reductions in mortality rates have been reported in the last couple of decades among preterm infants, there has been no reduction in the overall incidence of low birth weight and preterm births over the past several decades. In fact, the incidence of underweight births has risen steadily in recent years in many developed countries, such as France, Finland, and the United States. In Canada, recent statistics show that the low birth weight rate rose from 5.7% in 2001 to 6.1% in 2006, and the preterm birth rate increased from 6.4% in 1981 to 7.9% in 2004 (Canadian Institute for Health Information, 2007).

The consistent growth in the rates of low birth weight and preterm births may be attributed to a rising incidence of multiple births, greater use of assisted reproduction techniques, increases in the proportion of births among older women and greater use of elective Caesarean sections. It may also be due to the continuing increases in social inequalities between the rich and the poor, or between the socially included and the socially excluded. As mentioned above, the inequalities of adverse birth outcomes are observed both between and within countries. Although there are much higher adverse birth outcome rates in developing countries, low birth weight rates and preterm births also vary

significantly among socially developed societies. Low birth weight rates range from 4.4% in Finland (Moore, 1998) to 8.2% in the United States (Martin et al., 2007). Preterm birth rates are also quite different between Europe (6.2%) and North America (10.6%) (Beck, et al., 2010). Despite technological advances and overall wealthier conditions, preterm and LBW birth rates are higher in the United States (US) compared to other developed countries. The higher average income in the US does not lead to healthier average conditions and better reproductive outcomes of the population. This may be explained by its larger scale of income differences, increased social gaps between socio-economic classes or social positions and unequal access to health care.

For regions or larger areas, such as provinces or states, inequalities in adverse birth outcomes are persistently observed, even in countries with universal access to health care such as Canada. Low birth weight (LBW) rates ranged from 4.1% to 7% among Canadian provinces and territories in 2006-2007 (Canadian Institute for Health Information, 2007). Further down the spatial hierarchy, the inequality of low birth weight and preterm birth rates continues at smaller spatial areas among Canadian cities and among different neighbourhoods. Research has identified an association between neighbourhood incomes and LBW in the city of Toronto (Urquia, et al., 2007), British Columbia (Luo, et al., 2004), and Quebec (Luo, et al., 2006). These social and spatial inequalities of adverse birth outcomes need to be addressed carefully.

Because of their important impacts on people's health, adverse birth outcomes have been extensively studied and a great number of associated risk factors have been identified. These risk factors follow the framework defined by well-known social ecological models (Jeannine, 2010), which have in common the notion of multilevel systems of mutual influence and interaction on health outcomes, moving from personal level risks to larger social determinants including family influence, neighbourhood and community level risks, physical environmental risks, social institutions, and impacts from the state or even global systems. Despite these findings and established frameworks, public health interventions and practices still focus largely on personal or proximal risk factors. No effective health interventions or policies have been implemented to address macro-level social and environmental determinants of adverse birth outcomes.

This may be because it is usually difficult to intervene in and modify macro-level social determinants and risks. Interventions at higher levels of the social hierarchy usually are costly and involve the change of government legislation and regulations, social structures and norms. They also need the collaboration of multiple stakeholders from different sectors, not just the health institutions. The constant changes of the social system and status due to internal and external influences also require health interventions at these levels to be adjusted and adapted to the changes. Despite these difficulties, it is worthwhile to intervene in macro-level determinants given the many benefits this

may bring. The focuses of interventions at the macro level are on health promotion and disease prevention rather than treatment. Therefore, these interventions can target a much larger population compared with person-based health education programs. The targeted population includes the whole population or communities at risk rather than specific high-risk individuals. All members of communities may benefit by the interventions, not only in terms of their health, but also their quality of lives. High-level interventions also have impacts on all lower-level risks and may solve problems caused by social inequalities of health, which may not be able to be addressed by personal interventions.

## **1.1 Research gaps**

The focus of such macro-level interventions on primary prevention to pre-empt the onset of disease, illness, or other health problems is not to say that secondary prevention and tertiary prevention are less important. Following the frameworks suggested by the social ecological models, interventions are more likely to be successful when all levels of influence and their interconnections are addressed and primary, secondary and tertiary preventions are employed simultaneously. However, current research has only recently established the associations between various macro levels of risks and adverse birth outcomes, namely the direct associations between different exposure levels and adverse birth outcomes. While this may identify the targets of interventions at each level, there is a lack of supporting evidence for the interconnections between different level risk factors, and therefore a lack of knowledge as to how community-level health interventions can be approached. The indirect impacts and the pathways of the social and physical environment through individual risk factors on the outcomes of live births need to be examined to fill the gap in understanding of causality and the incidence of adverse birth outcomes.

Despite the fact that many individual and contextual risk factors have been identified, the etiological and social and spatial causes of adverse birth outcomes are substantially unknown. Personal-level risk factors may explain less than half of the total adverse birth outcome incidence in developed societies. There are even fewer acknowledged explanations for preterm births, most cases of which occur without any known causes. Among the already detected risk factors, some still show mixed evidence, such as the association between preterm births and maternal anxiety and depression caused by stressful life events, physical abuse and low levels of social support (Hoffman and Hatch, 1996). Some risk factors, such as those connected with micronutrients, have not been established statistically, due to methodological problems. Future research is needed to identify other etiological causes of low birth weight and preterm births and confirm the precise correlations between established risks and outcomes.



There is also insufficient knowledge of the exact causal pathways through which social determinants exert their impacts on adverse birth outcomes. It is still not clear how adverse social and environmental conditions are transferred to individual risk factors and thereby affect birth outcomes, and to what extent individual and contextual factors affect the inequalities of birth outcomes. Only a few studies have considered the role(s) that place and space play in determining low birth weight and affecting its various risk factors. This lack of knowledge poses a challenge for setting priorities and developing appropriate public health intervention programs and policies, especially for community-level interventions and interventions from the non-health sectors. The following section, bearing these questions in mind, proposes an overarching research goal and specific objectives to study the causal pathways of adverse birth outcomes and the spatial impacts through which social determinants may affect birth outcomes.

In addition, for community or neighbourhood level interventions, it is also necessary to target the groups or communities of focus for the interventions to be effective. The social and spatial interaction processes have led to an unequal spatial distribution of socio-economic and environmental risks and unequal risks of diseases, which have been demonstrated at different spatial levels. Using appropriate spatial analysis, the “hot spots” of significantly high risk areas of adverse birth outcomes can be identified and the unequal impacts of various social and environmental risks between different areas may also be addressed. This provides useful evidence to inform health professionals and health planners to conduct community-based health interventions to reduce adverse birth outcomes and to ameliorate their inequities.

## **1.2 Research goals and objectives**

The complexity of the risk factors of adverse birth outcomes and the difficulty of obtaining good data make it impossible to consider all of the potential risk factors in a single research project. Actually, the magnitude of the impact of each risk factor or determinant on adverse birth outcomes is different from place to place and from time to time. Hence, the purpose of this research is not to identify personal or medical risk factors for adverse birth outcomes. The upstream pathways from health to socio-economic status, as identified later in the materialist model, are also not the focus of this thesis. Rather, the purpose of this research is, by reviewing the current research on adverse birth outcomes, to identify and test hypothesized theoretical pathway models that may explain the causes of adverse birth outcomes and their inequalities. Of special interest is to examine the impact of space and place on their incidence (especially low birth weight and preterm births), or on the individual determinants of low birth weight, so that modifiable environmental or contextual risk factors can be identified to assist community-based and area-oriented interventions in order to reduce adverse birth

outcomes and improve personal health.

This thesis therefore seeks to answer the fundamental question, how and to what extent do social and environmental determinants affect the inequality of adverse birth outcomes? To be more specific, how and to what extent do unequal spatial distributions of social and environmental determinants affect inequalities in adverse birth outcomes in a socially advanced society? As most of the current research focuses on “what” types of social, environmental and individual risks may affect low birth weight and preterm births, the focus of this research is on understanding “how” the spatial distribution of social and environmental risks may affect adverse birth outcomes and their inequalities. If the roles of space and place in this relationship can be clarified, this may move the literature forward and help to improve current policies to reduce adverse birth outcomes, not only in the study areas but also in general, through targeted intervention strategies.

The goal of the research therefore is to clarify the role that place and environment play in the causal pathways of social inequalities that lead to birth outcome inequalities in a developed society, and at the same time to identify possible area-based and community-oriented interventions or policies to reduce these inequalities and reduce adverse birth outcomes. The specific objectives are:

- To define and test effective models that can best describe the impacts of personal, social and spatial determinants on adverse birth outcomes, and with available data, to use these models to describe both personal level and neighbourhood or community level determinants;
- To test and establish separate pathways that may cause different birth outcomes (preterm births, and LBW), especially the pathways that travel downward from neighbourhood-level risks to personal internal and external risks and eventually to adverse birth outcomes;
- To identify the direct and indirect roles that place and environment play on the courses and outcomes of pregnancy and on the social determinants of these outcomes;
- To determine how the social and built environments of different places (social integration, living conditions, accessibility to health care, socio-economic status, overall health conditions etc.) may influence differences in personal and behavioural risks, and eventually result in differences in adverse birth outcomes; and
- To suggest community-oriented interventions and possible improvements of current prenatal care policies to reduce the spatial and social inequalities of adverse birth outcomes.

The definitions of communities and neighbourhoods differ, in that communities may have a broader scope referring to both place (or locality) and identity (such as ethnic community or the gay community for example) (Campbell and Murray, 2004). However, the focus of this thesis research is on community as place. Thus, the two terms, community and neighbourhood, are used interchangeably for convenience in this thesis. Both of them refer to spatial areas that represent

individuals' places of residence or the surrounding environments where people live and interact with each other. People consider certain areas as their neighbourhood or their community that provide safe places and serves as important source of support and sociability (Wellman and Leighton, 1979). Within their neighbourhood, residents continually involve in neighbourhood affairs, activities, and interactions through community-level organizations and day-to-day communications. Through neighbourhood influences, common characteristics may be found that affect residents' behaviours or beliefs and consequently affect their health. These geographically associated characteristics may provide directions for community or neighbourhood-based health interventions to address population health issues.

When spatial impacts are mentioned in this thesis, the words “space” and “place” are frequently used. While both of these words denote common experiences about where we live, they actually differ from each other. In urban planning and health geography, “space” is more geographically or naturally defined representing background locations of where things are, which may be measured by distance, direction, area, and perhaps natural barriers. “Place” is more socially defined. Once a space has an identity, it turns to a place. People have a sense of a place by its associated identity, not by its location. For example, each built environment may have its associated physical, social, and environmental characteristics, and can be recognized and differentiated from neighbouring environments based on these characteristics. Neighbourhoods affect population health not only through geographical distance and natural barriers, but also through social distance and social barriers. Thus, in this thesis, the two words “space and place” are used together to represent potential spatial impacts on adverse birth outcomes.

### **1.3 Organization of the thesis**

The thesis is organized into six chapters. As indicated above, this chapter has introduced the background and the overall status of current adverse birth outcome research and practice, and proposed the goal and objectives of the research. Chapter 2 discusses in detail the identified medical, individual, organizational, social and environmental risks based on a framework of social ecological models. Chapter 3 introduces the general theoretical approaches of causal pathways of health inequalities and geographical perspectives of potential spatial impacts on health. Five hypotheses to be tested in the thesis are derived from this discussion. Chapter 4 provides the quantitative research framework and statistical models to transfer the hypothesized pathway models into statistical models and test them. Spatial analysis methods are also proposed to identify the spatial clustering of adverse birth outcomes and spatial impacts on different risk factors. Using the methods proposed in Chapter 4, a case study is conducted within three public health units in Ontario, namely the Halton Public Health

Unit, the Wellington-Dufferin-Guelph Public Health Unit and the Windsor and Essex County Public Health Unit. Chapter 5 describes the study settings, the individual and community-level data sources, and the statistical analyses. Based on the results, recommendations on community-based intervention and health planning are suggested. Finally, in Chapter 6, conclusions are made on the contributions of the research to the current literature on adverse birth outcome studies and methodologies. The generalizability of the research results and developed statistical models is discussed. Future research on adverse birth outcome studies is also recommended.

## **Chapter 2**

### **Determinants of Adverse Birth Outcomes**

Generally speaking, in most developed societies, pregnancy complications are few and outcomes are favourable for both mothers and infants. The most severe adverse outcomes of pregnancy are stillbirths or the death of the mother. Maternal death has become an extremely rare event in the developed world and stillbirths are also greatly reduced. However, even if both the mother and infant survive, perinatal pregnancy complications may lead to severe maternal or infant morbidity, both short- and long-term, especially for the early development and life-long health conditions of babies. This chapter, through a review of the literature, introduces the determinants of one of the most influential adverse birth outcomes, low birth weight (LBW). The two causes of LBW, namely preterm birth and intrauterine growth restriction (IUGR), and their identified various levels of determinants are introduced based on a social ecological health model. This review provides a solid academic background for the proposed research.

#### **2.1 Definitions and organizing framework**

The most commonly studied and most important determinant of perinatal, neonatal, and post-neonatal outcome is birth weight. Low birth weight is conventionally defined as a weight less than 2,500 grams for a live-born infant at birth. Low birth weight results from either preterm birth (defined as any birth that occurred before 37 complete weeks of gestation) (Arbuckle and Sherman, 1989) or intrauterine growth restriction or retardation (often represented by “small for gestational age” or SGA in short), or a combination of the two. IUGR and preterm births represent a significant health problem worldwide. While the major cause of LBW in developing countries is due to IUGR, the leading cause of LBW in developed societies is preterm birth. In Canada, approximately 70 percent of the LBW rate is attributed to preterm births. Although preterm births are responsible for only about 7% of all live births, they contribute to 75% to 85% of all neonatal deaths and about 50% of the long-term neurological impairment in children (Alexander, 1998; Joseph, 1998).

Researchers increasingly acknowledge that preterm birth and IUGR have different pathways to LBW and that each has distinct determinants (Paneth, 1995; Lang, Lieberman, and Cohen, 1996). The fundamental point for differentiating the causes of low birth weight is that preterm birth and IUGR may reflect different intra-uterine and perinatal processes and experiences, each of which may have specific implications for foetal development, neonatal morbidity and later childhood functioning.

While extreme IUGR is associated with neonatal death and short-term metabolic derangements, less extreme IUGR can cause long-term deficits in neuro-cognitive performance and growth (Hack, 1998; Goldenberg et al., 1998).

Extreme preterm birth (<28 weeks) is associated with high rates of mortality and of severe ophthalmological, pulmonary, and neuro-cognitive problems, chronic respiratory conditions, visual deficit, and infections (Dollfus et al., 1990). In addition, children born preterm are more likely to require special education and their impaired cognitive ability may result in poor performance and significant behavioural problems (Hack et al., 1995). It is therefore important to separate preterm birth and IUGR when examining the relationships of LBW with its personal and social determinants. Different causal pathways and risk factors can then be identified to address the problems of preterm birth and IUGR respectively.

Numerous studies have been published relating potential risk factors to either IUGR or preterm birth. The causes of IUGR are far better understood than the causes of preterm birth, although they share many of the same risk factors. Various risk factors or determinants are recognized at different levels of social hierarchies, both micro and macro, which can be better explained by a social ecological health model (Jeannine, 2010). As shown in Figure 2.1, this model organizes determinants of health according to five hierarchical levels of influence, namely intrapersonal, interpersonal, organizational, community, and society. The model recognizes the interwoven relationships between people and their living and working environments. While individuals are responsible for instituting

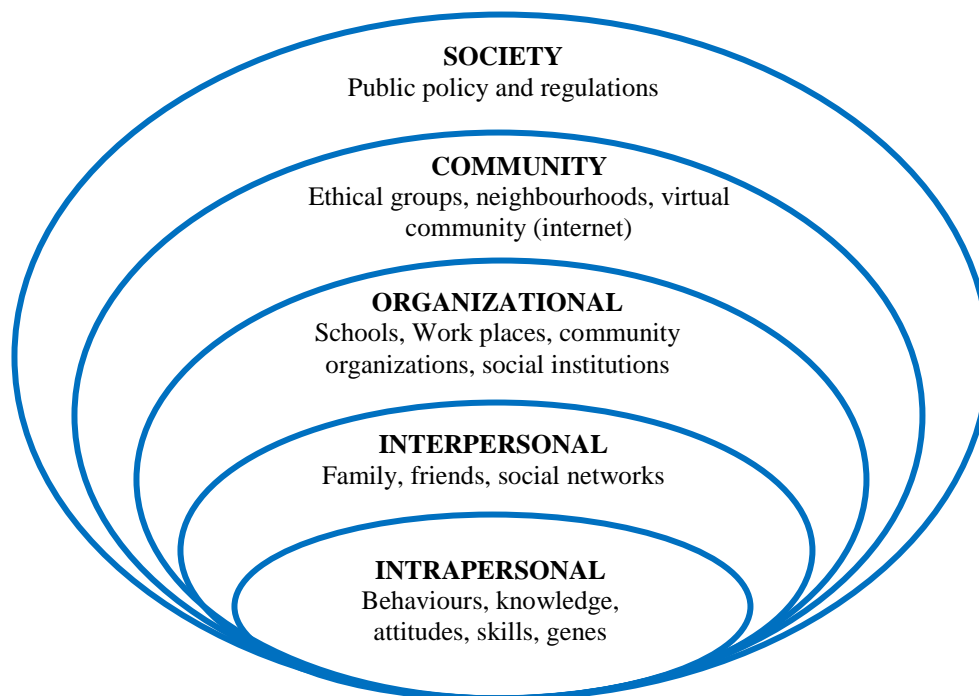


Figure 2.1: Social ecological model

and maintaining the lifestyles necessary for their health, personal behaviours and their health consequences are determined to a large extent by their social and physical environment, such as living conditions, community norms and values, working environment, health services, regulations, and policies. Barriers to health improvement are therefore shared among neighbourhoods, organizations, communities, or societies as a whole. Removing or lowering barriers at these macro-levels will make personal health behavioural change and health improvement more achievable and sustainable.

Factors influencing adverse birth outcomes at the intrapersonal or individual level are among the most often studied. These factors include biological and psychological factors, such as maternal demographic and anthropometric factors, genetic factors, maternal medical factors, maternal trauma, nutritional factors, infections, multiple births, stress, and lifestyles. At the interpersonal-level, factors such as family violence, family socioeconomic factors, and marriage status can be identified. The organizational level influences might include work conditions and work stresses. At the level of community, factors such as ethnicity, socio-economic class, social capital, social segregation, public facilities, the built environment, and environmental pollution may all have impacts. Finally, the most macro-level societal factors may include policy issues, infrastructure, and economics, such as early detection and secondary prevention policies and programs. The following discussion of adverse birth outcome determinants is organized based on these hierarchies.

## **2.2 Intrapersonal risks**

The intrapersonal or individual risks of adverse birth outcomes include both biological and psychological factors, internal physical condition and external health behaviours. The possible biological mechanisms are complex and multiple, and centre on the placenta, the foetus, the mother, and combinations of the three.

Biological factors that may cause IUGR in the womb are largely related to an abnormally small or blocked placenta, causing insufficient nutrients to reach the foetus, or other factors that prevent normal circulation across the placenta, causing poor nutrient and oxygen supply to the foetus (such as maternal under-nutrition, malaria, anaemia, actual and chronic infections) (Hendrix and Berghella, 2008). Other factors not involving the placenta include genetic or chromosomal anomalies of the foetus; low gestational weight gain due to inadequate energy intake, primiparity, multiple gestation, maternal age, unwanted pregnancy, or maternal anthropometry (such as low pre-pregnancy body mass index (BMI) and short maternal stature); congenital anomalies and genetic factors (such as familial proclivities for IUGR); maternal disorders (such as renal diseases and pregnancy induced hypertension); and pre-eclampsia (Kramer, et al., 2000).

The principal pathways leading to preterm birth are spontaneous preterm labour, preterm

premature rupture of the membranes (PROM), and medical induction (or iatrogenic preterm birth) (Berkowitz and Papiernik, 1993). Although the processes leading to preterm labour and preterm PROM are expected to have distinct epidemiological characteristics, there are relatively few studies that have assessed risk factors specific to either preterm labour or preterm PROM. Biological factors that may be related to preterm birth include genitourinary tract colonization, infection and inflammation, (such as bacterial vaginosis), multiple births (twin or higher order), maternal anthropometry (such as low BMI), maternal age, pregnancy induced hypertension, incompetent cervix, history of prior preterm birth and abruptio placenta. Micro-nutrients, such as vitamins, iron, zinc, calcium, and folate, may also be associated with IUGR and preterm birth, although the etiological role of micronutrients in either IUGR or preterm birth has not been confirmed statistically.

These factors are now discussed in detail in relation to IUGR and preterm births.

### **2.2.1 Maternal age**

The age at which a woman becomes pregnant impacts her likelihood of having adverse birth outcomes. Both teenage pregnancy and advanced maternal age (35 and over) can increase the risk of IUGR and preterm births.

Due to physical immaturity, teenage mothers may have reduced blood supply to the cervix and uterus and nutrient supply to the developing foetus, and an increased incidence of infections (Scholl, Hediger, and Schall, 1996). Externally, the psycho-social immaturity and low self-control of teenage women may also lead to higher incidence of unplanned pregnancies than in adult women (Elfenbein and Felice, 2003). The pregnancy may be concealed until a late stage due to fear, hence teenage mothers may consequently have delayed initiation of prenatal care (Scholl, Hediger, and Schall, 1996). Experiencing the pressures of crossing social boundaries, teenage mothers may also indulge in risky behaviours such as drinking alcohol, smoking, substance abuse, and resistance to a recommended nutritious diet (Roth et al., 1998). The interplay of these factors puts teenage mothers at higher risk than adult mothers to have adverse birth outcomes.

At the other extreme, although higher socioeconomic status, lower smoking rates, and healthier lifestyles associated with mothers of advanced age may have some positive impacts on IUGR and preterm births, older mothers still experience higher age-related chronic health problems, such as diabetes and hypertension, and higher pregnancy-associated complications, all of which can influence IUGR and preterm births (Newburn-Cook and Onyskiw, 2005). The reduced fertility potential and increased use of artificial reproductive technologies of advanced age mothers may also contribute to higher preterm and LBW births than for younger mothers (Tarlatzis and Zepiridis, 2003).



### **2.2.2 Maternal Anthropometric factors**

Three anthropometric factors may be associated with LBW and preterm births, namely maternal height, pre-pregnancy weight, and gestational weight gain.

An individual's height is a result of both genetic factors and development influences following birth, including environmental effects and nutrition. Studies have shown that short maternal stature may be associated with reduced birth weight (Prasad and Al-Taher, 2002; Blumenfeld et al., 2006) and preterm births (Honest et al., 2005). However, the etiological causes are not clear. Measured by BMI, which compares a person's weight and height to estimate a healthy body weight, both low BMI (underweight) and high BMI (overweight) are associated with increased risks of adverse birth outcomes (Honest et al., 2005). Mothers with low pre-pregnancy weight may have smaller infants as a result of life-long inadequacy of nutrition (Kirchengast et al., 1998). On the other hand, obesity may increase the risks of LBW and preterm births due to metabolic alterations, such as hyperglycemia and diabetes, and mild transient hypertension in pregnancy (Watkins et al., 2003; Bodnar et al., 2007). However, low BMI and short maternal stature are not always considered "risk" factors since the low birth weight in these cases may not represent true growth restriction or retardation in a pathological sense. They may merely reflect constitutional differences with no adverse consequences for the foetus or infants in their later development. Whether or not these factors may increase the risks of infant morbidity or mortality needs to be addressed in future research.

Inadequate weight gain of the mother during pregnancy is also associated with IUGR and preterm births. Maternal weight gain reflects increases in mammary and uterine tissues, fat stores, extracellular fluid and blood volume, and the development and growth of the foetus, placenta, and amniotic fluid (Johnson and Yancey, 1996). It is a marker for many physiological processes and conditions, as well as for nutritional status. Poor weight gain may be a reflection of a deficiency of substrates required for foetus growth (Carmichael and Abrams, 1997). It may also be due to marginal deficiencies, in particular micronutrients or inadequate macronutrient intake, such as zinc deficiency, which can cause appetite suppression, impaired synthesis of prostaglandins and collagen, uterine contractility, and finally lead to preterm births (Carmichael and Abrams, 1997). Poor weight gain may also be due to inadequate intake of macronutrients or energy. Malnutrition, especially in the first trimester, can result in poor plasma volume expansion and insufficient development of maternal tissues for support of the foetus (Carmichael and Abrams, 1997), which may lead to IUGR or preterm births. Therefore, it is necessary to maintain appropriate but not excessive gestational weight gain to optimize infant birth weight (Butte et al., 2005).

### **2.2.3 Genetic factors**

The biological role that genes play in the formation of IUGR or preterm births is currently not well understood. However, some epidemiological evidence suggests that genes may play a role in the system of determinants of adverse birth outcomes.

Birth weight is higher (about 150g on average) for male neonates than for female neonates. The difference is believed to be because of the effects of androgen, maternal foetal antigen difference, or genetic material carried on the Y (Amory et al., 2004). A mother or father who was born with LBW is likely to result in a LBW infant due to certain genetic conditions (Magnus et al., 1993; Klebanoff et al., 1998). The variability of birth weight distributions among different population groups is also believed to be due to maternal environmental and hereditary risks. For preterm births, the associated potential genetic predisposition may be due to a combination of factors responsible for a history of preterm births, racial predispositions, and the implication of certain single-gene disorders (Dizon-Townson, 2001). Overall, genetic factors are likely to play an important role in IUGR and preterm births. Further research into genetic variants associated with IUGR and preterm births, and the causative mechanism of the interplay between genetic factors and environmental impacts is needed.

### **2.2.4 Maternal health and medical conditions**

Maternal general health and chronic medical conditions, maternal infections, as well as pregnancy-associated conditions due to altered hemodynamic status can affect the foetus in several ways.

Maternal chronic conditions that affect nutrients and oxygen needed for foetal growth are key factors for IUGR and preterm births. The growth of the foetus and the duration of gestation can be affected by maternal chronic medical conditions that alter the intrauterine environment by affecting oxygen-carrying capacity, uteroplacental blood flow, and the size of the uterus (Alkalay, 1998). Chronic maternal hypertension can reduce foetal growth by a reduction in blood flow or the development of pre-eclampsia (Bernstein and Divon, 1997). Although the majority of infants born to mothers with diabetes are large for gestational age, maternal diabetes can still cause IUGR by longstanding changes in the microvasculature of the placenta (Bernstein and Divon, 1997). Hypoxemia caused by asthma attacks during pregnancy may affect either IUGR or preterm births (Murphy et al., 2006). Moderate to severe renal insufficiency is also reported to be associated with IUGR and preterm births (Ramin et al., 2006). Other maternal chronic conditions such as collagen vascular disorders, cystic fibrosis, inflammatory bowel diseases, starvation, pancreatitis, short bowel syndrome, malabsorptive states, cyanotic heart disease, and sickle cell anemia, may also have effects on foetal growth (Kliegman and Das, 2002; Pollack and Divon, 2002). If the mother has previously

had an abortion, this may lead to increased risk of preterm births due to infection (Sturchler et al., 1986) and mechanical trauma to the cervix during surgical procedures (Molin, 1993).

Gestational diabetes and hypertension are the two common pregnancy associated conditions that affect foetal growth. Pregnancy-induced hypertension may cause uteroplacental insufficiency and placental infarcts, which may restrict the growth of the foetus and induce preterm births (Villar et al., 2006). Superimposed gestational diabetes can cause IUGR if the mother has previous glucose intolerance (Bartha, 2003). In addition, maternal infection with organisms transmitted through the placenta during pregnancy, such as rubella virus, cytomegalovirus, malaria, syphilis, varicella, herpes, *Listeria*, and Epstein-Barr virus, can result in IUGR (Alkalay, 1998; Bernstein and Divon, 1997; Murphy et al., 2006). Other risk factors that may affect IUGR and preterm births include structural abnormalities in the uterus (Neerhof, 1995), placenta size and function, and the toxic effect of drugs, such as direct effects of antimetabolites, heroin, methadone, and alcohol on cell replication amphetamine, or the effect of cocaine and alcohol on the transport of amino acids (Ramsay and Goldenberg, 2002).

#### **2.2.5 Multiple births**

The increasing incidence of multiple births in most developed countries due to the advance of artificial reproductive techniques is one of the leading causes of preterm births and IUGR. The most common form of multiple births is twins although there are triplets or higher-order multiples. Multiples are likely to be growth restricted due to limited space in the uterus and a competition for nutrients among them, especially at the latest stage of the pregnancy. Starting from the 30th week of pregnancy, twins grow about 50 to 80 grams less per week than singletons. The risk is even higher among triplets, who are on average at the 10th percentile of the weight of a singleton baby at the 38th week. (Valero De Bernabé 2004).

More than 50% of twins and nearly all higher-order pregnancies are born preterm (Institute of Health Economics, 2008). The preterm labour is probably initiated by stretching of the myometrium. It is often preceded by pre-labour rupture of the membranes or expedited births due to maternal or foetal complications. In vitro fertilization (IVF) births have an increased risk of LBW and preterm births compared with non-IVF multiple births (McDonald et al., 2005). The development of national regulations for limiting the maximum number of embryos may be needed to reduce the risk of birth defects and disabilities.

#### **2.2.6 Maternal under-nutrition**

Nutrition is a key factor for foetal growth and a primary goal for each pregnancy. Major

nutritional factors that may affect foetal growth are intakes of both macro- and micro-nutrients of the mother, and nutrient uptakes of the foetus, including nutrient supply to the uterus and placenta, nutrients transport across the placenta, foetal uptakes of the nutrients, and foetal regulation of the nutrients. Nutritional requirements increase during pregnancy to support foetal growth. Appropriateness of energy and protein intake during pregnancy may be crucial. Maternal under-nutrition, especially in the late gestation stage, causes a slowing of foetal growth and may result in IUGR (Kliegman, 1990). Malnutrition may also cause stress in the foetus and lead to preterm births. On the other hand, excessive weight gain during pregnancy may also cause adverse birth outcomes due to its association with pregnancy induced complications, such as hypertension.

Dietary supplements may also have some impacts on birth outcomes. Fish oil may improve placental blood flow, postpone the onset of parturition, protect against pre-eclampsia and pregnancy-induced hypertension and subsequently improve foetal growth (Olsen, 2000). Micro-nutrients, such as iron, calcium, magnesium, zinc, Folic acid, Vitamin A, vitamin C, vitamin B complex, and minerals have all been shown to improve or potentially improve physiological and psychological parameters and tend to reduce IUGR and preterm births (Keen et al., 2003; Haider and Bhutta, 2006; Black, 2001). However, further studies are needed to assess and confirm the impacts of micronutrients on birth outcomes.

Research has shown that the nutritional status of a pregnant woman is largely affected by her socioeconomic status, lifestyle behaviours, and stress (Luke, 1994). Maternal under-nutrition is less a problem in developed countries than in developing countries, where starvation may still be a problem for a large part of the population. Provision of a balanced, nutritious diet to pregnant women in developing countries may be effective to reduce IUGR births. However, it may not be as effective in developed societies due to the relatively small prevalence of maternal under-nutrition and the lesser effectiveness with which adequate or extra nutrition can be transferred to the foetus. At the same time, interventions aimed directly at the foetus have not been adequately studied, and no effective intervention is available.

### **2.2.7 Maternal infections**

Infection plays a more important role in the onset of preterm births than in IUGR births. About one third of preterm births are believed to be associated with infections (Institute of Health Economics, 2008). Biological evidence suggests that the interplay of multiple mechanisms, including bacterial, viral, mycoplasmal, and parasitic infections, affects the onset of preterm labour and subsequent preterm and IUGR births.

Genital tract infections and urinary tract infections (UTI) are common causes of the spontaneous

onset of preterm labour. Genital tract infection may affect the genital tract environment by replacing the normal flora by infectious species, release various chemicals, rupture foetal membranes, infect the chorioamniotic sac, and subsequently lead to preterm labour (Hillier, 1995). Pregnant women are more vulnerable to UTI during pregnancy (about 20% incidences among pregnant women) due to morphological and physiological urinary tract changes (Santos et al., 2002). UTI is associated with preterm labour and pre-labour rupture of the membranes, according to epidemiological studies, but the exact etiological contribution is not well known (Millar and Cox, 1997). Other infections including trichomoniasis (Cotch, 1997), gonorrhea, syphilis (Doroshenko, 2006), influenza (Uchide et al., 2006), and human immunodeficiency virus (HIV) infection (Brocklehurst and French, 1998), are associated with increased risk of LBW and preterm births. Malaria and helicobacter pylori infection in pregnant women may also affect foetal growth and lead to IUGR (Yartey, 2006; Eslick et al., 2002).

Women's infections during pregnancy can be symptomatic or asymptomatic. Asymptomatic bacteria harboured in the urinary tract may result in serious adverse birth outcomes if untreated. Thus, it is helpful for both mother and foetus to have routine serological screening for commonly encountered viral infections in the first trimester and early identification and treatment of bacterial infections.

### **2.2.8 Maternal stress and depression**

Maternal psychological factors, such as stress and depression, may also be associated with IUGR and preterm birth. Associations have been established between preterm birth and stressful life events, depression, a physically and mentally demanding job, and low level of social support (Copper et al., 1996; Nordentoft, et al. 1996; Collins et al., 1998). However, the results are mixed. Not all studies have proved a connection between preterm birth and stressful life events. In addition to dramatic stressful life events, chronic stressors may also be able to predict low birth weight instances. Although less dramatic, these chronic stressors, including daily coping issues, happen more frequently and accumulate during the mother's early life-course and may eventually affect birth outcomes.

The exact mechanisms of how stress may lead to the onset of IUGR and preterm births are still not known. However, there is growing evidence of an interaction of neuroendocrine and immunological processes in the initiation of adverse pregnancy outcomes (Wadhwa et al., 2001). Stresses may disturb a mother's health balance by repeated activations of the fight-or-flight response (Brunner, 1997), which may increase the concentration of glucocorticoids and catecholamines and cause cortisol-induced increases in placental secretion of a corticotrophin-releasing hormone (CRH). The release of CRH increases the production of prostanoids, which can stimulate uterine contractility

and induce preterm birth (Hobel et al., 1998). The activated fight-or-flight response may also disrupt the hypothalamic-pituitary-adrenal axis and trigger the onset of preterm births (Hogue, 2001; Chan, 1993). In addition, the release of stress hormones may lead to immunosuppression and alteration of both cellular and humoral immunity. The changed immune responses may make the mother susceptible to infection, which, as discussed earlier, can be responsible for IUGR and preterm births (Romero, 2001).

Stress may also increase the risk of preterm birth indirectly through the development of a depressive self-concept (Seguin et al., 1995; Metalsky et al., 1993) and low commitment to pregnancy (Lydon et al., 1996). The depressive self-concept includes dissatisfaction with one's current state, a sense of hopelessness, and a lack of optimism about the future. It may contribute to heightened stress appraisal in response to daily chronic stresses. This may create a downward circle whereby heightened stress appraisal and the depressive self-concept reinforce each other. The depressive self-concept may induce the release of stress hormones, such as the above mentioned placental CRH. It may also lead to behaviour changes that increase the risk of preterm birth and IUGR, such as irresponsible sexual practices and enhanced susceptibility to genital tract infection and inflammation, and low commitment or less purpose to have a baby. The later is also associated with irresponsible or careless health behaviours such as reduced intakes of nutrients, delayed prenatal care, increased use of cigarettes, cocaine and other drugs, and heavy alcohol drinking during pregnancy. All of these indicate that psycho-social factors may play an important role in causing adverse birth outcomes.

### **2.2.9 Life style and health behaviours**

In addition to the above internal biological and psycho-social factors, maternal external health behaviours and life styles also play important roles in adverse birth outcomes.

Cigarette smoking is one of the most important risk factors related to IUGR and preterm births. Smoking results in twice the likelihood of a low-birth weight delivery, an average reduction of 200g in birth weight (Walsh et al., 2001), and twice the likelihood of a stillbirth (Kesmodel et al., 2002). Smoking during pregnancy is recognized as the most important preventable risk factor for adverse birth outcomes and it has been largely reduced in recent years. Despite this, the strong association between smoking and adverse birth outcomes still makes smoking an important etiological cause.

The effect of smoking on birth weight has been shown to be dose related, that is, birth weight is reduced in proportion to the amount of tobacco smoked (Hebel et al., 1988). Even passive smoke or environmental smoke is significantly related to low birth weight through IUGR (Misra and Nguyen, 1999; Windham et al., 1999). Although the mechanism of the effect of smoking on the foetus is not completely established, among the thousands of compounds in tobacco smoke, nicotine, metabolite

cotinine, and carbon monoxide contribute the most to adverse pregnancy outcomes.

Nicotine can freely cross the placenta and concentrate in the foetal circulation and amniotic fluid at higher concentrations than in maternal circulation. Nicotine in the maternal blood is metabolized to cotinine, which has a much greater half-life than nicotine and therefore can reach much higher levels in the maternal plasma, and in the foetal circulation. The actions of nicotine decrease uterine artery blood flow, cause changes in umbilical artery flow and foetal oxygenation and acid-base balance, decrease foetal heart rate, and increase mean arterial pressure. Carbon monoxide (CO) also crosses the placenta rapidly and freely and concentrates in the foetal circulation, reducing the availability of oxygen to the foetus tissues and causing hypoxia (Andres and Day, 2000; Pastrakuljic et al., 1999). Other compounds from tobacco, such as cyanide, also compete with oxygen and lead to hypoxia.

All of these may restrict foetal growth and cause IUGR and preterm births. In addition, smoking may cause placental abruption by altering maternal prostacyclin (Busacca et al., 1984), and cause infection by changing immunological responses. Smoking may also alter the maternal nutrient environment by disturbing the amino acids and zinc transport (Kuhnert, et al., 1987). Due to altered tastes, smoking mothers have poorer quality diet, eat less, and have higher nutritional deficiency, which may increase the risk of rupture of foetal membranes and preterm labour (Zaren et al., 1997).

Heavy alcohol drinking is another important risk behaviour related to IUGR (Zuckerman and Hingson, 1986; Walpole, et al., 1990). It has been established statistically that binge drinking can be associated with a trend towards increased risk of IUGR and preterm births (Lundsberg et al., 1997). More than two drinks per day was associated with birth weight reduction by 200g and the effects are suggested to be dose responsive (Abel and Hannigan, 1995; Clarren, 1988). In addition, binge drinking may also increase the possibilities of unintended pregnancy and associated risks (Naimi et al., 2003).

Although the exact mechanism of alcohol-induced impacts on adverse birth outcomes is still not clear, some theories have been proposed to explain some of the biological phenomena. Alcohol can cross the placenta barrier freely and reach the foetus. The foetus is therefore exposed to equivalent levels of ethanol to the level in maternal circulation. It is, however, less efficient for the foetus to excrete alcohol than it is for the mother. This is especially true during the first half of pregnancy. The prolonged circulation of alcohol breakdown products, such as acetaldehyde, is fetotoxic (Lundsberg et al., 1997). Alcohol-induced risks to the foetus are related to high levels of blood alcohol concentrations. Mild alcohol intakes otherwise may have a protective effect due to a mild estrogen increase in the blood, which may protect the overall perinatal outcomes (Mariscal, et al., 2006). However, the threshold level is unknown.

Coffee and tea are widely consumed in the general population throughout the world. The

biological mechanisms of the effects of caffeine on pregnancy outcomes are not clear. Some biological evidence indicates that high caffeine consumption may have a negative impact on foetal growth and IUGR (Laudignon et al., 1991). However, the association is not well supported by epidemiological evidence (Browne, 2006; Leviton and Cowan, 2002). As well, there is no evidence of an effect of caffeine use on preterm births.

Cocaine use is another important modifiable determinant for adverse pregnancy outcomes since it is one of the most commonly used substances for recreational purposes, especially among populations from large metropolitan cities in North America (Holzman, 2001). Cocaine may affect the foetus through various mechanisms. Although the biological mechanism is still not clear, epidemiological evidence suggests increased risks of cocaine use for IUGR, preterm births, placental abruption, and prelabour rupture of the membranes (Addis, et al., 2001; Holzman and Paneth, 1994). Other substances used during pregnancy, such as amphetamines and methyl-amphetamines (Plessinger, 1998), and narcotics (Hulse, et al., 1997), are also associated with increase risks of IUGR and preterm births. The evidence of the effect of marijuana on IUGR and preterm births however is conflicting (Fried and Smith, 2001).

Exercise during pregnancy is an important question that needs to be addressed. Exercise could be beneficial for a number of reasons. Exercise may improve mothers' immunological defence mechanisms and prevent urinary tract infection, and therefore prevent the possible consequence of preterm labour. Improved muscle tone by exercise helps during labour. Exercise may also increase foetal weight by improving the mother's blood flow (Hatch et al., 1993). However, heavy exercise is also risky since it is stressful and there is a chance that vigorous leisure time aerobic exercise in the third trimester of pregnancy may invoke preterm labour (Hatch et al., 1998).

Delayed or reduced prenatal care, such as not making enough prenatal visits or starting visits at the second trimester or later, are also acknowledged as risk factors for adverse birth outcomes (Alexander and Korenbrot, 1995). The reasons that cause fewer and late prenatal visits are manifold. Violence in the surrounding neighbourhood may deter the mother-to-be from leaving her house. She may not have access to transportation, or the clinic may be too far to make the visit. She may not have the financial support necessary to take time off work to make the visit. She may be too young and not have sufficient education to understand the importance of neonatal instructions. She may also have an unwanted pregnancy and therefore be careless to maintain her pregnancy health. Many of these factors may in one way or another be linked to mothers' low socio-economic status and their adverse living environment. Thus, delayed or reduced prenatal care is not only a personal behaviour or personal responsibility, but also a social responsibility.

Overall, health behaviours and life styles are considered to be the most effective modifiable



determinants of adverse birth outcomes and are commonly addressed by health practitioners. While prevention of and attention to the associated factors show much success in reducing adverse birth outcomes, interventions in health behaviours or other personal-level risks are limited in their capacity to address risks caused by higher-level determinants, especially for addressing social inequalities of adverse birth outcomes. Higher level risks therefore need to be identified and addressed.

## **2.3 Interpersonal risks**

Good intimate relationships may stimulate creative and meaningful living, and provide safety and mutual understanding. Characteristics such as mutuality, intimacy, affection, sympathy, loving and ability to form deep lasting relationships are commonly identified as defining psycho-social health (Heath, 1983). On the other hand, adverse interpersonal relationships between family members or friends, or instances of social isolation, have negative medical consequences and increased risk of illness and health. Adverse pregnancy outcomes are one of the many health consequences that may result from these factors.

### **2.3.1 Marital status**

Probably the closest interpersonal relation is marriage. Marriage is reported to have a positive effect on the general health of the partners (Raatikainen et al., 2005) and is supposed to affect pregnancy outcomes positively as well. Marriage provides a stable relationship and protects mothers through the practice of healthy behaviours and positive attitudes. Married mothers are better off in these respects compared with unmarried mothers, many of whom are younger, unemployed, and have lower socio-economic status, higher stresses related to birth, and riskier behaviours such as smoking, and heavy drinking (Waldro, Hughe, and Brooks, 1996). Compared to the unstable relationships of cohabitant or common-law unions, marriage may also increase the likelihood of social, psychological, emotional, and financial support provided by the partner in reducing stress and its consequences on adverse birth outcomes (Raatikainen, Heiskanen, and Heinonen, 2005).

### **2.3.2 Family violence or abuse**

Family violence or abuse against women at any time in their lives represents serious social, legal and medical problems. Violence during pregnancy is even more harmful since it poses particular threats to the mother and additional threats to her foetus. Studies report the prevalence of intimate partner violence to be between 3.9% and 8.3% (Gazmararian et al., 1996). The incidence is higher in teenagers and adolescents (23% to 37%) (Cherniak, et al., 2005), and in aboriginal populations compared with non-aboriginals (Heaman, 2005).

Physical abuse can affect pregnancy through both direct and indirect mechanisms. Violence may

cause direct trauma to the mother's abdomen, which leads to release of arachidonic acid and initiates contractions and preterm labour and delivery (Pak et al., 1998). Physical abuse may also increase the risks of miscarriage, ruptures of foetal membranes, placental abruption, and foetal fractures. Other rare adverse consequences may include rupture of the uterus, liver or spleen, and pelvic fractures (Sammons, 1981; Murphy, 2001).

A significant indirect influence of violence and abuse is the result of psychological stress, which may lead to maternal depression and adoption of risky or dangerous behaviours during pregnancy such as smoking, excessive alcohol use, illicit drug use, careless food intake, or inadequate use of health services (Covingto et al., 2001), any of which consequently lead to IUGR and preterm births. Verbal abuse about the mother's weight during pregnancy may lead to purposeful weight control and may predispose to IUGR (Campbell et al., 1999). Unintended pregnancies due to abusive sexual relationships may also lead to irresponsible health behaviours during pregnancy and lead to IUGR or preterm births (Pallitto et al., 2005).

Interventions in screening and harm reduction have been recommended or implemented in practice. However, there is a lack of evidence for or against screening for violence by healthcare professionals in a regular setting during the prenatal period (Ramsay et al., 2002; MacMillan and Wathen, 2003). Interventions to reduce harm from violence during pregnancy have indicated some degree of benefit for mothers (Wathen and MacMillan, 2003). However, the effects of such interventions in reducing adverse birth outcomes are still not clear.

## **2.4 Organizational risks**

Organizations, such as companies, schools, professional associations and community organizations, are systems with a formal multi-echelon decision process operating in pursuit of specific objectives (Miller, 1978). Individuals are probably connected to many different organizations in their lifetime and are affected or constrained by the physical conditions, social or cultural environments, rules or norms of the organizations, which may have consequent impacts on their health outcomes.

### **2.4.1 Occupational Factors**

The effects of occupation on pregnancy outcome could be due to hard physical activities, stresses associated with job activities, or exposure to various substances at work. For substance exposures, the biological mechanisms that may cause adverse birth outcomes could be different for different substances, and it is not feasible to discuss in here the associations of many work-related substance exposures and their impacts on birth outcomes. The impacts of some of the environmental substance

exposures on birth outcomes are discussed later in the environmental pollution subsection.

For work-related physical conditions, evidence from epidemiological studies shows that duration of work, type of work, and workplace activities all influence birth outcomes. Hard, physically demanding work that may require prolonged standing or strenuous activities may lead to IUGR and preterm births. Although the biological mechanisms underlying the effect are still not well understood, it is suggested that several mechanisms may be simultaneously at work. Heavy strenuous work that involves prolonged standing can lead to increased sympathetic vasomotor tone to skeletal muscles, leading to compromised uteroplacental perfusion and diminished nutrient and oxygen supply to the foetus (Simpson, 1993). Large uteroplacental infarcts may be induced for mothers who still work in a standing position at a late stage of gestation (Naeye and Peters, 1982). Excessive work activities may also result in hyperthermia, which may have adverse effects on the foetus (Simpson, 1993).

Highly demanding jobs with low self-control or less feeling of success may also lead to stress and depression, which is, as discussed earlier, an important contributor to IUGR and preterm births. Some women continue to hide their pregnancy for fear of losing their jobs and even continue to work strenuously. This creates both physical and psycho-social burdens to their domestic responsibilities and may lead to adverse pregnancy outcomes. Certain blue-collar jobs are associated with a higher risk of preterm births and low birth weight births compared with white-collar jobs (Chia, et al., 2004; Virji and Talbott, 1990).

The influences of other organizations, such as schools, professional associations, and community organizations, on adverse birth outcomes have not yet been examined based on an exhaustive review of the related literature.

## **2.5 Community and society level risks**

Community and societal level risks combined with organizational risk are often called macro-level risks or macro-level determinants in order to differentiate them from micro-level medical and behavioural risks. These potential risks or determinants related to IUGR and preterm birth include racial/ethnic origin, socio-economic status (income, education, and employment), environmental pollution, low access to medical care and prenatal care, poor living conditions, low access to local health related services (such as food groceries and pharmacies), lack of social support, stressful working environment, inadequate health policies related to prenatal care, adverse community characteristics, such as low levels of walkability, insecurity, lack of social integration, and lack of community-based health intervention or prenatal care.

Although racial/ethnic origin and socio-economic status are seemingly personal characteristics, they are really factors that cross many different levels in the social hierarchy. They are rather to be

conceptualized as macro-level determinants instead of personal characteristics because the disadvantages of being a member of a minority and of low socio-economic status are usually common for members within each population or subgroup due to culture, economic factors, environment, and behaviours associated with these groups. In other words, the disadvantages of being a member of a minority or being poor are not merely determined by personal traits, but, more importantly, by upper level social structures and society-level determinants. Major community and society level risks are discussed as follows.

### **2.5.1 Environmental pollution**

Exposure to environmental toxins can be harmful for birth outcomes. Although biological mechanisms are not well understood for most environmental toxins, some epidemiological studies support an effect of pollutant exposures on adverse birth outcomes. Due to the complex composition of air pollution, the accuracy of different methods of measuring exposures, and the different interpretations of these factors in the literature, no consistent conclusions have been made and no effective intervention policy has yet been developed. Research is still needed from a local perspective for policy purposes.

Air pollution is one of the major concerns. Air pollution varies from place to place. The highest levels of air pollution occur near industrialized areas or in the path of air-borne pollution sources. Among the many airborne pollutants, total suspended particles and some gases (including sulphur dioxide, carbon monoxide and nitrogen dioxide), are the major substances implicated in air pollution (Bobak, 2000). Although the biological mechanisms for the effects of air pollution on adverse birth outcomes are not well understood, epidemiological associations are established between suspended particulate and gaseous air pollutants and adverse birth outcomes (Ritz et al. 2000, 2002; Wang et al. 1997; Xu et al. 1995; Bobak 2000; Dejmek et al. 1999; Bukowski, 2004).

Exposure to air pollutants leads to increased risks of maternal infection and illness (Gibbs, 1992). Air pollutants are absorbed into the maternal bloodstream and lead to an increase in blood viscosity and may affect placental blood flow, or may cross the placental barrier and have direct toxic effects on the foetus. Some other air pollutants, such as benzopyrene, have anti-estrogenic effects, which may affect uterine health and lead to IUGR (Bui et al., 1986). Studies suggest that the foetus is most susceptible to the effects of air pollution during the first trimester (Generoso et al. 1987; Rutledge 2000). Initial changes due to air pollution might be triggered in early pregnancy, around the time of implantation (Khong et al. 1986), and lead to IUGR.

For other environmental pollutants, a weak association has been observed between exposure to high levels of disinfectant by-products of drinking water and IUGR (Bove et al., 2002). Exposure to

pesticides during pregnancy may lead to preterm births, but the estimated impact on LBW or IUGR is inconsistent among studies, since identification of an individual chemical as the cause could be difficult due to the presence of various chemicals in pesticides (Hanke and Jurewicz, 2004). Some observational studies also suggest that excessive noise exposure may influence IUGR and preterm births, and may cause high-frequency hearing loss and minor congenital malformations of the foetus (Nurminen, 1995; American Academy of Pediatrics, 1997). There are also indications from epidemiological studies that seasonal variation may have a role in affecting birth weight (Lawlor et al., 2005; Elter et al., 2004). However, the exact biological mechanisms of how and to what extent environmental pollution may affect adverse birth outcomes are still not clear. Further studies are needed.

### **2.5.2 Race/ethnicity (culture vs. personality)**

Although social and racial differences are among the most extensively studied social determinants, neither the exact mechanism nor the interventions to alleviate their adverse impacts on adverse birth outcomes are known. In countries which have multi-racial/ethnic populations, inequalities in pregnancy outcomes are generally found among groups with different racial and ethnic origins. Higher IUGR rates have been reported among blacks in the United States, Asians in the United Kingdom, aborigines in Australia, and both Negros and Asians in the Netherlands. In Canada, a higher risk of low birth weight is found among recent immigrants (Urquia et al., 2007), although a lower rate of preterm birth is also associated with these groups.

The variability of birth weight distribution among different racial groups is suggested to be due to different racial predispositions to maternal environmental and hereditary risks (Dizon-Townson, 2001). However, while genetic differences may play some role in influencing adverse birth outcomes, it has also been posited that race itself is not really the explanatory factor. Most of the inequalities are linked to socio-economic disadvantages among racial minority groups, and the interplay of stress and racism (defined as racial discrimination and prejudice). Minority racial or ethnic groups may experience racism at different levels, including personal, institutional, cultural, and collective racism (Utsey, 2002). Residential or institutional segregation between minority and majority ethnic groups may be formed where people of minority groups are socially excluded. They are sometimes unfairly treated, have limited access to services and opportunities, and are sometimes perceived to be of lower socio-economic class. Mothers of minority members who experience these kinds of prejudice or discrimination may internalize them, and develop feelings of inferiority, loss of control, frustration, or become depressed. Racism does become an important stressor that causes psycho-social problems and leads to preterm birth and IUGR. In the United States, studies show that black mothers living in

hyper-segregated metropolitan areas have higher risks of preterm births than black mothers in non-hyper-segregated metropolitan areas. Black-white preterm birth disparities were also larger in hyper-segregated areas than in non-hyper-segregated areas (Osypuk and Acevedo-Garcia, 2008), reflecting potential psychosocial and environmental hazards associated with population-level patterns of racial and ethnic inequality.

Accumulated cultural changes due to long-term exposure to different social or ethnic environments may also be responsible for related health problems. The formation of common norms or life styles during cultural changes may sometimes be negative. For example, the adoption of unhealthy group behaviours, such as tobacco use, alcohol intake, and illicit drug use, will have negative impacts on adverse birth outcomes.

### **2.5.3 Socio-economic status**

Socio-economic inequalities in pregnancy outcomes, such as LBW, IUGR, preterm birth, and infant morbidity and mortality, are among the most robust findings in perinatal epidemiological research. It is not surprising to find such inequalities in countries like the United States, where there are large differences between the rich and the poor. However, this situation is also true in countries like Canada, Finland, Sweden, Scotland, and Spain, where there are lesser degrees of poverty and there is universal access to high-quality prenatal care (Moore, 1998; Martin et al., 2007; Beck, et al., 2010; Canadian Institute for Health Information, 2007).

Evidence shows that preterm birth and IUGR are consistently more frequent among the socially disadvantaged groups in almost all developed societies (Wilkins et al., 1986; Luo et al., 2006). Low pre-pregnancy BMI, reduced weight gain during pregnancy, reduced nutrient intake, increased tobacco, alcohol, coffee, and illicit drug use, stressful work and living environment, delayed or reduced prenatal care, increased maternal infections, violence and abuse, depression, increased risk of unwanted pregnancy, increased teenage pregnancy, and reduced levels of social and financial support were all found among mothers of low socioeconomic status (Kramer et al., 2000).

Socio-economic inequalities do not have direct impacts on adverse birth outcomes. Poverty and income inequalities are considered as antecedent to the above mentioned personal-level exposures and behaviours. In other words, socio-economic disadvantages operate at the upper level. They lead subsequently to unhealthy behaviours, exposure to stress and psychological reactions to stress (from work or residential environment), poor living conditions, and environmental risks that may increase the risk of IUGR or preterm birth. Evidence has been found that within less socially advanced groups, people tend to be short, have low gestational weight gain, low intakes of micronutrients, more frequent and heavier smoking, more common uses of cocaine, marijuana, and narcotics, high alcohol

and caffeine consumption, inadequate initiation and frequency of prenatal care, more strenuous work, more stressful life events, depression, and lower level of social support.

However, the interaction between socio-economic status and physical, behavioural, and psychosocial risk factors does not follow a simple route. Socio-economic status operates not only downstream, but also upstream in that unhealthy behaviours, depression, and poor living environments may lead to low education, unemployment and low income. This feedback forms a downward cycle for socially excluded groups, who suffer not only from generally worse health, but also from lower quality of life.

It is easy to argue that the way to bring about a reduction of the impact of socio-economic inequalities on health or on adverse birth outcomes in particular is to reduce socio-economic inequalities. However, with the rapid globalization of the world's economy, socio-economic disparities are not likely to be reduced. Indeed, income inequalities have increased during the last three decades in most of the developed societies (Morone and Jacobs, 2005). If this situation is not going to be solved in the near future, it is important to understand the causal pathways through which the effects of social inequalities travel down the social hierarchy from broad structures to local geographical areas, and from macro-determinants to micro- personal level risks. This will assist the evolution of alternative interventions that can be addressed to reduce the adverse impact of socio-economic inequalities on health.

#### **2.5.4 Neighbourhood conditions**

One important route through which socio-economic inequalities affect population health in general, and adverse birth outcomes in particular, is the condition of space and place. The social divisions of different socio-economic classes lead to spatial divisions within urban areas, especially within and between residential neighbourhoods. Socially disadvantaged groups are usually clustered in segregated urban areas (Meng et al., 2005). These areas are separated from residential areas of the mainstream classes and are often socially deprived. These neighbourhoods are more likely to be exposed to environmental risks, such as pollution (air, water, noise, radiation, and chemicals), have poor living conditions and constrained social interactions, have poor access to healthy foods and be deprived of medical facilities, available green spaces, walkable environments and good quality schools. These factors not only lead to unhealthy conditions for the local residents, but also constrain their opportunities to learn and to thrive, which in turn may result in poor health.

A long history of research has shown that health status (e.g., mortality, morbidity, birth weight) and other aspects of individual well-being and behaviours vary strongly across neighbourhoods (Kawachi and Berkman, 2003). Recent studies have established correlations between adverse birth

outcomes and numerous neighbourhood characteristics, including physical, socio-economic, and psycho-social factors. Some studies accessing neighbourhood-level disadvantaged environments, deprivation, and cumulative exposure to income inequality have clearly shown that neighbourhood-level factors are significantly associated with infant birth weights (Reagan and Salsberry, 2005; Sastry and Hussey, 2003; Buka et al., 2003; Farley et al., 2006). In some neighbourhoods, adverse effects of racism or racial stigma may increase the risks of adverse birth outcomes for minority women despite socioeconomic advantages (Pickett et al., 2005).

Social deprivation (measured by community economic hardship, housing costs, unemployment rates, percentage of low education and so on) is associated with low birth weight deliveries (Pattenden et al., 1999). Incidences of teenage pregnancies are much higher in high-poverty neighbourhoods (Harding, 2003). Adverse neighbourhood physical environments, such as proximity to landfill sites (Rushton, 2003), are also harmful for a mother's reproductive health. Pregnant women living in proximity to supermarkets that sell fresh produce and other healthy food have significantly fewer low birth weight births than other pregnant women regardless of income level (Lane et al., 2008). Women living in urban areas may also be more likely to have IUGR or term SGA babies compared to women living in suburban locations (Elo et al., 2009).

Neighbourhood mechanisms related to stress and adaptation (violent crime, reciprocal exchange, and participation in local voluntary associations) are reported to be one of the most robust neighbourhood-level predictors of birth weight (Morenoff, 2003). Mothers residing in neighbourhoods with favourable perceived security have lower risks of SGA birth than neighbourhoods with less favourable perceived security (Auger et al. 2008). Neighbourhood factors also interact with personal risk factors to affect adverse birth outcomes. Personal-level risk factors, such as adverse health behaviours, operate differently depending on the characteristics of the neighbourhood. For example, although prenatal care is universally available, women living in high-risk neighbourhoods may benefit less from prenatal care than do women living in lower risk neighbourhoods (O'Campo, et al., 2007).

#### **2.5.5 Prenatal health care provision and education**

Due to the well established importance of adverse birth outcomes to babies' early development and life-long health conditions, both primary and secondary prenatal care interventions have been provided to most pregnant women throughout the world in order to reduce adverse birth outcomes. Primary care mainly focuses on the prevention of the onset of adverse birth outcomes by health advice and early education, psychosocial support, monitoring and early diagnosis, while secondary prenatal care provides treatments to threatened IUGR or preterm labour through medical interventions.



Prenatal health advice and education are currently delivered through clinic visits, prenatal classes, telephone help lines and outreach health workers. These forms of intervention are mostly person focused and many of them focus on at-risk women only. The content of education includes suggestions of general care plans, basic elements of investigations during pregnancy, suggestions of nutritional supplementation, information on signs and symptoms of adverse birth outcomes, advice on clinical visits for investigations of infections, and brief teachings regarding risk factors, such as social issues, health, lifestyle, and family history of disease and adverse birth outcomes (Lumley and Donohue, 2006; Sprague, 2004). Intensive interventions for smoking cessation through counselling, multiple contacts, supportive materials and follow-up have shown to be beneficial in reducing the rate of LBW births. However, educational programs and prenatal care directed towards high-risk women do not show preventive effects on preterm births (Hueston et al., 1995). A possible reason may be that programs that educate pregnant women in the early recognition of preterm labour may generate additional anxiety in pregnant women from information overload, which may be detrimental. There was, however, a trend of reduced risks of LBW births for women provided with baseline information and a booklet about prevention of adverse birth outcomes, especially for a subgroup of black women over nineteen years of age (Moore et al., 1998).

Observational studies have demonstrated some beneficial effects of psychological support. Current methods of providing social support vary. Commonly practiced methods include home visits by health professionals, such as midwives, social workers, or nurses, or specially trained laypersons at regular intervals to provide individual counselling and tangible assistance on psychosocial issues, social support provided by the mother's own social network, and group prenatal care provisions. Studies show that programs that offer individual support during pregnancy as part of clinical prenatal care intervention are unlikely to prevent IUGR and preterm births (Hodnett and Fredericks, 2003; Blondel and Bréart, 1995). Group prenatal care facilitated by midwives or obstetricians in some programs has been found to provide benefits to reduce LBW (Ickovics et al., 2003). The success of group prenatal care programs is believed to be due to the content of psycho-social prenatal care, higher intensity of prenatal care in the group setting, changes in social norms, avoidance of risky behaviours, opportunity to discuss these behavioural changes, and increased contact time. Proper identification of groups experiencing higher chronic stress during pregnancy and provision of psychosocial support to those at-risk groups or communities may be justified.

Screening and early diagnosis are commonly provided through ultrasonography, recording of blood pressure, measurement of weight gain, urinalysis, blood tests, and various other methods (White, et al., 2002). Different risk factor-based scoring systems using factors such as maternal medical and social history, social and behavioural risk factors, results of a rapid fibronectin assay, and

sonographic measurement of cervical length, have been shown to identify at-risk populations (Honest et al., 2004; Creasy et al., 1980; Tekesin et al., 2005). The combination of history, physical examination, biochemical tools, and sonography is capable of identifying at-risk pregnancies and provide borderline predictive accuracy (Creasy et al., 1980; Tekesin et al., 2005), although further research is still needed for better predictions. As discussed earlier, chronic stressors may cause the increase of corticotropin-releasing hormone (CRH). Despite the fact that the exact role that CRH plays in the early detection of preterm labours is not well understood, studies on the ability of CRH to predict preterm labours are promising (Ramsay and Goldenberg, 2002; Leung et al., 1999). This provides some evidence to support the role that psycho-social factors play relative to the onset of adverse birth outcomes. However, the detection of adverse birth outcome risks does not necessarily lead to their reduction. Detection should be supplemented by interventions or measures that prevent adverse outcomes.

Once a mother is diagnosed with threatened preterm labour or fetal growth restrictions, secondary preventions need to be put into place to prevent the onset of adverse birth outcomes, improve foetal maturation before birth and reduce the side-effects and risks to the mother. Various interventions for threatened preterm labour have been practiced, including bed rest, hospitalization, hydration, tocolytics, cervical cerclage, and antibiotics. Evidence suggests that administration of antibiotics for infection during pregnancy is beneficial for women with preterm PROM, but not to women with threatened preterm labour with intact membranes (Andrews et al. 2006; Morency and Bujold, 2007; Kenyon et al., 2004). However, it is a clinical dilemma regarding whether to treat the mother or not because antibiotics may be effective to the mother but may not be as good for the fetus, and may cause some deleterious consequences (Kenyon et al., 2003). Some tocolytics, such as betamimetics, and progestational agents (progesterone) may reduce the incidence of IUGR and preterm births in high-risk pregnancies (Anotayanonth et al., 2004; Coomarasamy et al. 2006). However, evidence shows that bed rest, hospitalization, and hydration are not effective to reduce the onset of preterm labour (Sosa et al., 2004; Crowther, 2001; Goulet et al., 2001; Yost et al., 2005; Stan et al., 2002).

Multi-component programs are also undertaken in various settings, including education, social support, screening and medical treatments (Papiernik, and Goffinet, 2004, Armson et al., 2001; Armson et al., 2003; Scott et al., 2001). These programs are mostly directed towards high-risk pregnant women. The results of these programs are variable for reducing the incidence of adverse birth outcomes and most of them are unsuccessful. It is argued that the problem may be due to the fact that most of the interventions in these programs have not been shown to be effective in preventing preterm births when each of them is used as an isolated intervention (Institute of Health

Economics, 2008).

Prenatal health care interventions provide certain mechanisms to alleviate the impact of socio-economic inequalities on adverse birth outcomes. However, the provision and consumption of health care are not always equally distributed among individuals due to various social, political, and geographical factors. Health service provision and distribution is influenced by political and administrative structures. The decisions of where, how, and to what extent to provide health-related services involve a trade off between different stakeholders and different social classes to preserve existing political structures, more than they reflect public need.

Historical reasons and financial situations of the local governments and other organizations also affect the performance of different health intervention programs. These programs or services may not be equally available, accessible and affordable by every mother-to-be (Penchansky and Thomas, 1981). They may be arranged in a fashion that favours groups who possess power, which is especially harmful for the poor because they are the least powerful group. Geographical constraints also influence health care usage. The spatial distance to health care services and the means of transportation limit accessibility to health services, especially for remote rural area residents and individuals without cars and with restricted physical mobility.

#### **2.5.6 Social capital**

The other possible buffering mechanism on the impact of socio-economic inequalities and related stressors is the increase of social capital within neighbourhoods or communities. The general interrelations of social capital and health are discussed more extensively in Chapter 3. Although the exact mechanisms of social capital on adverse birth outcome are currently not well understood, discussions earlier in this chapter have shown the impacts of psycho-social factors and the benefits of social support and group-level interventions on adverse birth outcomes.

In places where neighbourhoods have high social capital and local residents are more engaged in the social life of their community, residents are more likely to generate informal resources by assisting one another with favours, providing each other with health-related advice and other information; aiding one another with everyday tasks, monitoring each others' property, and participating in local voluntary associations within their community. This may provide not only a safer, liveable, and mutually beneficial living environment, but also a mechanism to buffer the impacts of various stressors from outside lives, which may eventually improve people's health in general and reduce women's adverse pregnancy outcomes in particular. Through this mechanism, pregnant women may therefore achieve better access to informal resources through social relationships and community institutions to adapt to adverse environments, stressful life events and

daily chronic stressors, and to get first-aid quickly when needed. All of these are beneficial for improving their birth outcomes.

Social capital may be achieved by careful developments of community-oriented policies or community-based programs to encourage public and multiple stakeholder participation, increase social interaction and integration, improve community health services and other facilities, and build mutual trust among neighbours.

## **2.6 Summary**

Overall, from an individual point of view, a woman is more likely than others to have a IUGR or preterm birth if she has had a LBW or preterm baby before, she is too young or too old, she is carrying more than one baby, she is having a first baby, she smokes during pregnancy, she is a heavy drinker, she is underweight, she is not getting enough healthy food, she has a lot of stress in her life or is depressed, she has a vaginal or bladder infection, she has a chronic health condition or has a serious medical problem, she has uterine or cervical anomalies, she has had miscarriages before, she does strenuous work during pregnancy, she is a single mother, or she is poor. As discussed earlier, current public health and hospital interventions focus mainly on these individual risks to prevent adverse birth outcomes. These kinds of interventions include early prenatal care, early prenatal education, and advice on healthy behaviour such as smoking cessation or cut-down, avoiding strenuous physical activity, healthy eating, stress management, social support, screening and earlier detection, and medical treatments. While some of the interventions are successful in reducing adverse birth outcomes, the person-oriented interventions fail to address higher-level risks and may not be able to reduce the inequalities caused by these social and environmental risks.

Community health intervention or promotion programs provide a mechanism to address macro-level determinants of adverse birth outcomes. Such programs may only be achieved by multiple stakeholder involvement and inter-sector co-operation among local government departments, public health units, hospitals, planning departments, local at-risk residents, and local community organizations. Indeed, the idea of community-oriented interventions is consistent with prevalent planning concepts, such as “new urbanism”, “smart growth”, and “sustainable development”.

The main principles of these concepts that may be related to health include raising residential densities, providing more mixed land uses and pedestrian-friendly environments, encouraging public transit, revitalizing old existing neighbourhoods, creating more affordable housing, adopting more diverse regulations concerning aesthetics, street layouts and design, increasing social capital and social resilience, creating a more socially and economically integrated environment and so on. However, although some community-oriented program strategies have been planned (Best Start,

2002), these principles are more talked about than actually implemented in current health planning practice due to various historical, political, and economic reasons (Downs, 2005). The establishment of the connections between these planning principles and health improvement may promote collaboration actually to implement them. Both health and planning may benefit by reducing environmental and financial cost, improving population health, and generally enhancing peoples' quality of life.

The failure of many multi-component preterm birth prevention programs shows that the lack of knowledge of the effect of each individual component on reducing adverse birth outcomes may be an important obstacle to success. The lack of evidence on the mechanisms of the pathways of social and environmental determinants in general and the unknown local spatial variations of identified risk factors in particular impede the effective implementation of community or neighbourhood oriented interventions to reduce adverse birth outcomes. The next chapter discusses some generally accepted theoretical approaches that seek to explain health inequalities and the roles that space and place play in these processes. From this discussion, the research hypotheses are stated for this thesis regarding adverse birth outcome inequalities.

## Chapter 3

# Theoretical Approaches to Studying Inequalities of Health and Adverse Birth Outcomes

The research discussed in Chapter 2 showed that adverse birth outcomes are associated with various social and environmental determinants including socio-economic conditions, ethnicity, living conditions, environmental pollution, health-related services and psycho-social conditions. To be able to investigate these risks needs a clear understanding of their effects and operating mechanisms or pathways. As non-communicable health outcomes, adverse birth outcomes follow most of the general theories and approaches relating to the causal pathways and geographical perspectives on health inequalities, especially approaches appropriate to chronic diseases and general health conditions. By understanding these general theories and approaches, plausible hypotheses on the pathways of community- or neighbourhood-level risks toward adverse birth outcomes may be identified, and the impacts of neighbourhood settings or spatial arrangements on adverse birth outcomes may also be addressed. Thus, the purpose of this chapter is, through a general discussion of health inequality theories, to derive testable hypotheses on inequalities of adverse birth outcomes, which are the focus of the quantitative analysis in this thesis.

### 3.1 Theoretical models of health inequalities

The original conceptualization for explaining the relationships between health and the environment is the “agent-host-environment” triad (May, 1958), which is used to explain the occurrence and spread of infectious diseases (Figure 3.1).

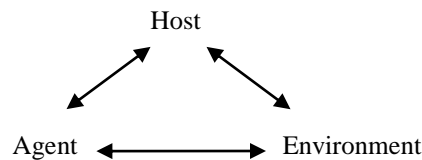


Figure 3.1: Agent-host-environment Triad (Adapted from Young, 2005, p115)

The diffusion of infectious diseases is modeled based on this triangulation. For example, tuberculosis requires an agent (the *Mycobacterium tuberculosis*), a host (the patient who may be a malnourished alcoholic), and an environment (such as an overcrowded shelter shared with others who are infected). This conceptualization is based in “germ theory”, or the biomedical model, which

believes that microorganisms are the causes of disease (Susser, 1985). Health practises based on this theoretical approach focus on disease protection of individuals through regulation and standards to control the more direct, biophysical effects on personal health. They are successful in overcoming infectious diseases.

Once non-communicable diseases became dominant in a society, the less obvious pathogenic causal pathways of chronic diseases lead to a shift in the object of study from microorganisms to the social determinants of health. Researchers from a variety of disciplines other than medicine, including geography, planning, ecology, psychology, demography, anthropology and sociology, have attempted to explain the causes of health inequalities from a macro-level systematic view of health. Many social models have been established to explain the possible causal pathways of health inequalities. The social ecological model described in Chapter 2 is one such model suitable in framing the overall hierarchies of health determinants and explaining the interrelations of different levels of risks and their influences on personal health outcomes. However, the model lacks the ability to explain individual pathways through which social determinants may lead to different outcomes. The significance of individual pathways to health intervention planning and implementations makes it necessary to understand separate mechanisms that may lead to different health results for different social groups.

Four widely accepted pathway models can be distinguished in the literature that attempts to explain the causal pathways of health inequalities, namely the behavioural model, the psycho-social model, the material/neo-material model, and the life course approach. These models form the basis for understanding health inequalities in most developed countries. Each model is discussed in turn.

### **3.1.1 The behavioural model**

Health behaviours are intimately linked to social inequalities. The health survey for England shows that the less advantaged an individual's social class position, the more likely the person is to smoke, be regularly drunk, or be obese (Marmot and Wilkinson, 2006). Behavioural models emphasize that health inequalities are determined by personal behavioural choices (Townsend et al., 1986, Townsend et al., 1992). These models therefore seek to explain why behaviours might vary systematically between different social groups.

A “direct behavioural explanation” assumes that unhealthy behaviours are mostly caused by adverse personal characteristics (Osler et al., 2000). This explanation argues that people who are less endowed with certain types of personal characteristics, such as intelligence or coping skills, have less control over their circumstances and are therefore more likely to have bad health behaviours. It implies that genetics play an important role in explaining the variation of behaviours among groups of

people – people behave differently because of their different genetic compositions. Based on this concern, genetic composition is often suggested as an explanatory factor in the assessment of variations in disease risks, although alternative explanations that favour a stronger environmental and social influence may be more plausible.

The “cultural explanation”, on the other hand, emphasizes the effects of social position or circumstances on health behaviours (Blaxter, 1990). It does not view risky behaviours or social disadvantages as consequences of personal characteristics. Rather, it argues that culture, as a system of values and meanings, serves as a standard of behaviour and defines the way of life peculiar to a society or a segment of it, where similar norms, values, knowledge, and belief are shared (Helman, 2000). Culture may affect people’s belief in a healthy life, educational achievement (Lynch et al., 1997), self-regulation in maintaining central social role (Siegrist, 2000), and expression of distinction from others (Bourdieu, 1984). These effects may in one way or another affect health behaviours and cause health inequalities.

Combining the direct and culture explanations, the causal pathways of the behavioural model can be expressed in Figure 3.2. In this model, the focus is the behavioural impacts on health outcomes, although genetic composition may directly affect personal health and disease in some circumstances. The debate is whether behavioural differences are indigenous (caused by differences of individual personalities) or external (acquired through cultural differences among groups).

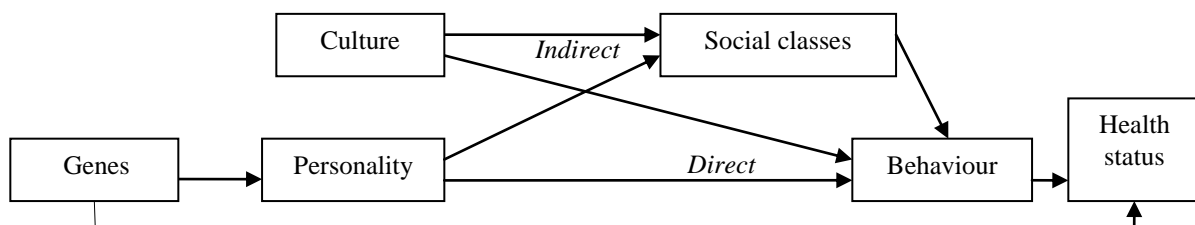


Figure 3.2: The causal pathways of the behavioural explanation on health inequalities

### 3.1.2 The psycho-social model

In most studies of health inequalities, the health behaviours discussed in Figure 3.2 have not been found to explain all of the social differences (Bucher and Ragland, 1995). One alternative explanation is the psycho-social model (Figure 3.3), which argues that relative position in the social hierarchy affects people’s feelings and imposes direct stress effects on a person’s health status (Elstad, 1998; Theorell, 2000). A number of psycho-social factors may cause stresses, such as lack of control and autonomy at work (Bosma, et al., 1998), the imbalance between home and work, and the imbalance between efforts and rewards (Peter and Siegrist, 1997). In addition, “buffering factors”, such as social support and social affiliation, are also important to moderate the ill effects of stress (Steptoe, 2000).



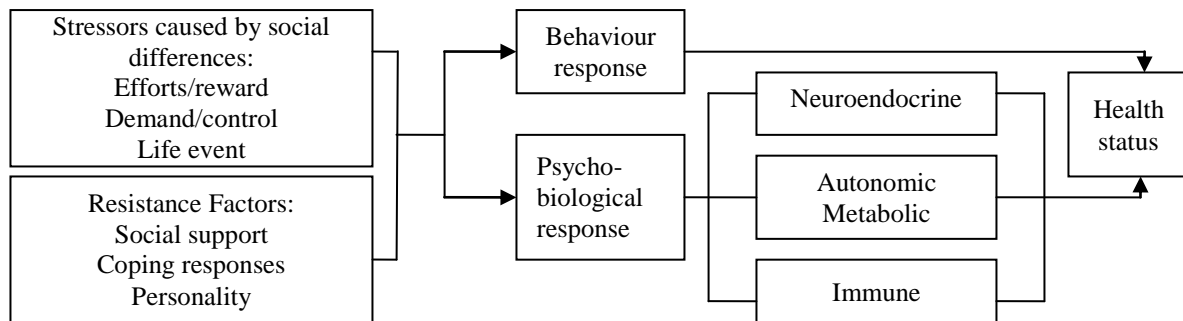


Figure 3.3: The causal pathways of the psycho-social model (Adapted from Brunner and Marmot, 2006, p27)

Stresses may disturb health balance by the repeated activation of the fight-or-flight response (Brunner, 1997), which involves two main pathways, namely the sympatho-adrenal pathway and the hypothalamic-pituitary-adrenal axis. These two pathways are known collectively as neuroendocrine pathways, which together co-ordinate an array of metabolic and physiological changes by the reaction of the nervous system and several hormonal or endocrine transmitters to stresses. The resulting physiological changes provide the means to survive in the face of emergent environmental challenges. However, in advanced industrialized countries, physical or biological emergencies are rare and people usually have adequate food, clean water, and sufficient material circumstances. Life is otherwise filled with psychological demands and challenges, such as financial strains, lack of control at work, lack of social support, and lack of sense of success. These subtle exposures to stresses may activate the fight-or-flight response too hard, too often and for too long, which may have multiple health costs and induce some chronic diseases, such as diabetes and cardiovascular disease. The possible roles that psycho-social factors or chronic stressors may play in producing adverse birth outcomes were discussed in Chapter 2.

Since people with lower social status and fewer resources have less possibility to control their situation and fewer chances of rewards, they tend to have more prolonged exposure to psychological demands and stresses, and therefore are more likely to have psychologically-related health problems. This explains the fine “social gradients” that such diseases increase progressively down the social strata (Steptoe, 2000).

### 3.1.3 The materialist and neo-materialist model

No matter how strong the psycho-social argument is, the differences in social positions are largely built on the material differences of social classes. In addition, material inequalities, such as income inequalities, in themselves determine the inequalities of people’s working conditions, living environment, access to health care facilities, and exposure to physical hazards, which may all have some impact on population health. Poor health is found in almost all studies of geographical areas

characterized by poor material conditions, such as poverty or pollution (Dunn, et al., 2006; Malmstrom et al., 1999; and Mackenbach et al., 1993).

The purely materialist model therefore suggests that the health variations of different social groups are determined by their varying degrees of exposure to risks or physical hazards in their living and working environment over the life course. These risks are part of the social structure, over which the individual has no control (Blane, et al., 1998). They can include consumption of unhealthy food, uncomfortable living conditions, dangerous working environment, high levels of pollution, and traffic and other accident dangers. Health status is determined by the accumulation of risk exposures. More socio-economically advantaged people have less chance of exposure to direct material risk factors than less advantaged people, and therefore they have fewer health risks. In addition to direct risk exposures, personal wealth and material conditions may also determine relative social position, which may put psycho-social stresses on individuals of lower social status and consequently generate health-threatening behavioural responses (Figure 3.3) (Jarvis and Wardle, 2006).

The more complex materialist explanation argues that the way that low income and power cause people to have poor housing is not the same as the way that low income and power cause lower-paid jobs or work hazards. The relationship of low income and housing takes place in the sphere of consumption (low income results in less buying power to have adequate housing). The relationship of income and employment conditions takes place in the sphere of production. A person cannot buy a job in the same way as a house. However, people with higher income and power may have higher quality of individual development, more qualifications or credentials, and more social interactions or influential contacts than people of lower income (Raphael, 2004). This, in turn, may give them better life chances and bargaining power in the labour market. They therefore have more chance to get a safer, cleaner, better-paid job. This may create an upward cycle for the economically advanced groups. On the other hand, a downward cycle may also be created for the disadvantaged groups, in which low incomes lead to fewer opportunities, lower-paid jobs, poorer health conditions, and, in turn, lower incomes.

In addition, the neo-materialist approach emphasizes that living conditions and employment are also determined by higher-level social-structural factors, such as public service provision, welfare and health coverage, and social and economic policies (Lynch et al., 2000). Social structures shape social classes, determine the distribution of resources over social and physical infrastructures, and consequently affect the quality of various social determinants of health.

Thus, combining both materialist and neo-materialist perspectives, a person's health situation is a result of the overlapping impacts of both social structures and living conditions (Figure 3.4). While there are common focuses between the materialist and neo-materialist explanations of the relationship

of physical conditions of individuals and their health, such as unemployment and poor housing, the concentration of interventions are different. The materialist explanations concentrate on what different incomes can buy relative to the health of individuals, such as food, transportation, services, living and working conditions. The neo-materialist perspective, on the other hand, concentrates on how public provision and the distribution of services and policies can affect population health inequalities.

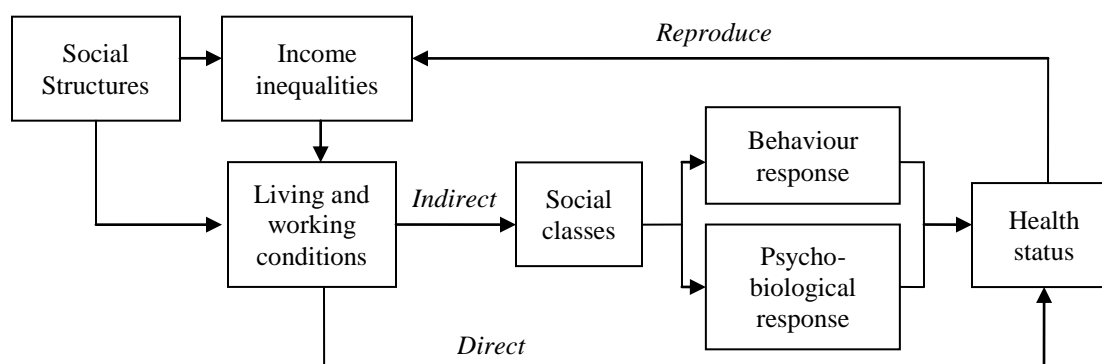


Figure 3.4: The causal pathways of the materialist/neo-materialist model

### 3.1.4 The life-course approach

All the above three models have a distinctly non-historical emphasis and focus on adult risk factors. However, evidence shows that growth and early development are closely associated with adult chronic disease (Kuh and Ben-Shilmo, 2004). A life-course approach then emphasizes the accumulated effects of experience across the life span in understanding the maintenance of health and the onset of disease (Blane, 2006; Dannefer, 2003). In this approach, adult health status is a result of complex combinations of physical, social and economic circumstances taking place over time, accumulated from their very early childhood and even before birth (Bartley, 2004).

Given the wide range of exposures to risk factors over the life course and their potential timing and duration, exposures may affect disease risks in a variety of ways. Several models can be identified for explaining the life-course impacts of social determinants of health (Ben-Shlomo and Davey Smith, 2002). The first is the “critical period model” or “latency model” (Hertzman et al., 2001). It argues that if a hazard or adverse experience takes place at a certain age, it may have lifelong detrimental effects on the functions of tissues, organs, and body systems that will not be reversed in any dramatic way by later experience. For example, certain risks, such as German measles, are dangerous at a certain time during gestation. An extension to this model also argues that the lasting effects of exposures in the early critical period may interact with exposures in later life either to enhance the effects on chronic diseases or diminish them.

The second model is the “accumulation of risk model”, which assumes that risks or advantages to health accumulate gradually over the life course (Kuh and Ben-Shlomo, 2004). In this model, the time that a risk is experienced is not important, but it is important whether other risks precede it or follow it. Risk exposures may be independent of each other (Figure 3.5a) or clustered (Figure 3.5b).

One plausible reason for clustering is that many risk factors are correlated with a person’s socio-economic position in society. For example, low childhood or family socio-economic status is associated with low birth weight, higher stresses, inadequate food, inadequate housing, lower quality of education, unpleasant living environment, and health-threatening behaviours, which together lead to later poor adult health.

Model 3.5(c) is the “chain of risk” model, which looks for interactions between risk factors. It argues that, in some cases, an exposure is only damaging to health in certain groups of people because of a sequence of previous exposures. The build-up of harm happens through a pathway – one bad experience or exposure leads to another and so on, which eventually damages people’s health (Hertzman, et al., 2001). An early exposure in a chain of risks may have a trigger effect only for the next exposure and it may be that only the final exposure in the chain has a marked effect on health (Figure 3.5c). For example, if a person is not exposed to the mumps virus in childhood and developed immunity, he may get serious disease as an adult. Alternatively, each exposure may also have an independent “additive effect” on later diseases in addition to increasing the risks of later exposures (Kuh and Ben-Shlomo, 2004) (Figure 3.5d).

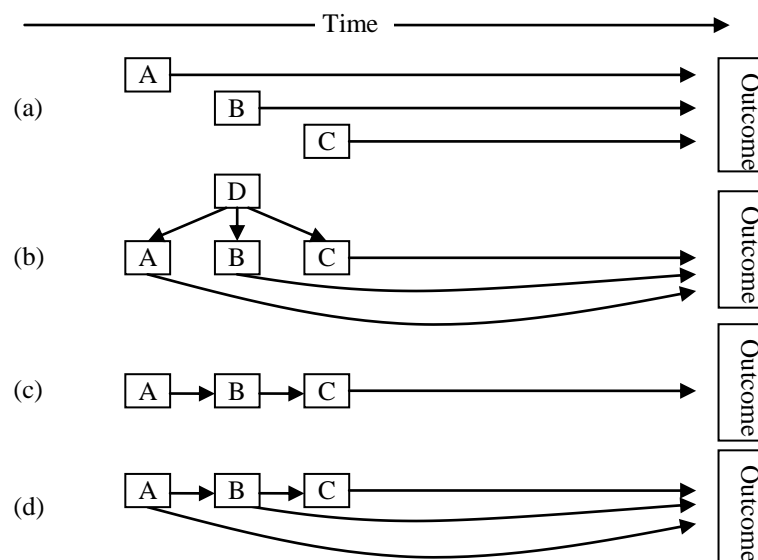


Figure 3.5: Life course models (where A, B, C, D denotes risk factors) (After Kuh and Ben-Shlomo, 2004, p10)

The life course models are not mutually exclusive and may operate simultaneously. For example,

the accumulation of risks does not preclude the possibility that factors acting at critical developmental periods have a greater impact than at other times. Risk factors can act independently, collectively, or in a chain of pathways over the life-span to affect health.

To summarize, the above four causal models, namely the behavioural model, the psycho-social model, the materialist/neo-materialist model and the life course approach, are different from each other. Each of them has its own influences, focuses on social aspects, and policy implementations. These differences are organized and compared in Table 3.1.

<i>Model</i>	<i>Influences</i>	<i>Focuses</i>	<i>Policy</i>
Behavioural	<ul style="list-style-type: none"> <li>• Direct behavioural explanation: personal behavioural choices, coping skills, intelligence, and life styles are responsible for their developing and dying from diseases.</li> <li>• Culture explanation: Differences in beliefs, norms and values cause members of social groups to have different educational and other achievements and have different health behaviours.</li> </ul>	<ul style="list-style-type: none"> <li>• Health behaviours, such as smoking, drinking, diet, exercises, leisure time activities and sexual activities.</li> </ul>	<ul style="list-style-type: none"> <li>• Changing behaviours through education and legislation.</li> </ul>
Psycho-Social	<ul style="list-style-type: none"> <li>• Direct impact: long-term stresses weaken the immune system, increase insulin resistance, and lead to greater incidence of lipid and clotting disorders and other biomedical injuries that serve as precursors of disease in adulthood.</li> <li>• Indirect impact: high level social inequalities cause stresses, which induce health-threatening behaviours and consequently lead to health inequalities.</li> </ul>	<ul style="list-style-type: none"> <li>• Increased social hierarchy, unfavourable self-social comparison, and weakened social cohesion and social networks.</li> </ul>	<ul style="list-style-type: none"> <li>• Increasing social capital or collective efficacy through participation, social network building, and neighbourhood design.</li> <li>• Improving working conditions.</li> </ul>
Materialist/Neo-materialist	<ul style="list-style-type: none"> <li>• Material conditions determine personal physical, developmental, educational, and social status; this status in turn affects material conditions.</li> <li>• Material conditions lead to differences in the experience of psycho-social stress, which have direct impact on human body.</li> <li>• Material deprivation and stress lead to health threatening behaviours.</li> <li>• Social structures determine the individual material conditions</li> </ul>	<ul style="list-style-type: none"> <li>• Material conditions: such as income inequalities, living environment, transportation, and public services.</li> <li>• Social and economic structures: resource distribution.</li> </ul>	<ul style="list-style-type: none"> <li>• Public intervention on service provision.</li> <li>• Improving working and living conditions.</li> <li>• Equal distribution of resources.</li> <li>• Market regulation.</li> <li>• Social welfare.</li> <li>• Environmental risk control.</li> </ul>
Life-course	<ul style="list-style-type: none"> <li>• Social determinants of health operate at every critical period of development – prenatal, early childhood, childhood, adolescence, and adulthood.</li> <li>• The social and environmental impacts are accumulated across the life span.</li> </ul>	<ul style="list-style-type: none"> <li>• Accumulated impact of social and physical .</li> <li>• environment risks on health across the life span.</li> </ul>	<ul style="list-style-type: none"> <li>• Early life intervention, improving parent and child relationships, and education.</li> <li>• Adult health promotion and elderly assistance.</li> </ul>

Table 3.1: Comparison of influences, focuses, and policy implementation of health inequality models

The differences evident in the models in explaining the pathways of health inequalities reveal that

each of them has its own advantages and disadvantages (Table 3.2). Each of them is good at explaining only a certain aspect of the social determinants of health. There is no one universally applicable explanation that excludes all others. It is far more likely that several causes are simultaneously at work.

<i>Model</i>	<i>Advantage</i>	<i>Disadvantage</i>
Behavioural	<ul style="list-style-type: none"> <li>• This model identifies the impacts of behaviours on health inequalities.</li> <li>• Personal characteristics and group cultures are linked to health behaviours.</li> </ul>	<ul style="list-style-type: none"> <li>• The social and environmental impacts on health behaviours are not well explained.</li> <li>• Health-related behaviours only explain part of the differences in health risks. Other factors which directly influence personal health are not addressed.</li> </ul>
Psycho-social	<ul style="list-style-type: none"> <li>• This model effectively explains the fine grain social gradients of health inequalities (the lower the social position the poorer the health) in most of the developed countries.</li> <li>• This model identifies the impact of psychosocial stress on health.</li> <li>• The importance of social support, social cohesion and social integration for health are identified in this model.</li> </ul>	<ul style="list-style-type: none"> <li>• Solely looking at the impacts of stresses and social capital ignores the impact of increasing material differences on perceived social position and feeling.</li> <li>• This model cannot address the direct health impact of material inequalities, such as income inequalities, which determine the inequalities of people's working conditions and living environment.</li> </ul>
Materialist/ Neo- Materialist	<ul style="list-style-type: none"> <li>• The direct material impacts on the human body, on personal health behaviours, and their feelings are addressed.</li> <li>• The impacts of social and political structures on resources distribution and then on health inequalities are identified.</li> <li>• Health inequalities induced by income or socio-economic status are well explained.</li> </ul>	<ul style="list-style-type: none"> <li>• Material differences are not sufficient to explain health differences between more and less privileged workers.</li> <li>• Income or material inequalities are not the only reasons for health inequalities between countries. Some more egalitarian states have higher health inequalities than other less egalitarian states.</li> <li>• This and the above two models are snapshots of adult health inequalities.</li> </ul>
Life-course	<ul style="list-style-type: none"> <li>• This model effectively captures early life impacts on health inequalities.</li> <li>• This model identifies the accumulated effects of social determinants on health inequalities, and therefore makes possible the combination of the above three explanations.</li> </ul>	<ul style="list-style-type: none"> <li>• This model is relatively new and therefore there is less evidence to support the argument.</li> <li>• The longitudinal view of this model and the integration of other causal pathways make it complex and hard to measure.</li> </ul>

Table 3.2: Advantages and disadvantages of health inequality models

Proximal causes may also operate differently in different societies, on different population groups, in different locations, and at different time periods, or interact with others to produce inequalities in health (Berridge and Blume, 2003). For example, in a highly privatised society in which income is unequally distributed, income level is a key marker of worth, power, status and identity. The aggregation of income inequalities in certain areas may become a significant factor for aggregated health inequalities. On the other hand, in a society in which there are more egalitarian redistributive tax policies, welfare, and universal health coverage, income may not be a significant marker for social stratification, which may be framed by other factors such as education or personal

achievement. The aggregation of inter-area income inequalities may have less influence on aggregated health inequalities. Thus, health inequalities are society- or context-oriented phenomena and should be understood based on the local context.

Despite their differences, the above four pathway models overlap in one way or another through the interplay of intimate material, behavioural, and psychological factors, each of which is constrained by higher level social, environmental and cultural structures. The interplay of all these factors eventually determines personal health. Different combinations of pathways provide the opportunities to integrate the causal models to explain the variation of health inequalities over different societies and different regions.

The life-course approach links all risk factors across a person's lifetime and therefore provides a possible way to integrate the other three causal models in explaining health inequalities. During the life-span, each additional advantage in a certain group in terms of material, psycho-social conditions or culture difference, experienced either by oneself in present time, in childhood, or by one's parents, may produce a higher chance of good health and long life. The accumulation of these advantages eventually determines health differences between social groups. A life-course neo-materialist would also argue that life time exposure to material or psychosocial conditions is also structured by social and economic policies and institutions (Bartley, 2004).

Figure 3.6 shows an integration of the above models into a single framework (Marmot and Wilkinson, 2006). It is important to qualify this framework by noting that it is only one of many possible integrations. While each model has its own distinct role, there are no clear boundaries between these models.

In general, a hierarchical structure can be identified in the framework with social structures and institutions on the top, neighbourhood and communities, living and working conditions, and social and cultural relationships in the middle, and personal behaviours, psycho-social stress, socio-economic status, genetic factors, and pathophysiological pathways at the bottom. However, the causes of diseases do not necessarily follow a simple chain that can be reduced to personal level factors. Rather, the plausible pathways are complex and it is the interplay of social factors with biological factors, physical environment, and the person's history that shapes a particular health manifestation (Parkes et al., 2003). In the next section, the debate over the impacts of contextual and compositional (personal) factors on health inequalities is examined in order to illuminate further the complex and hierarchical nature of the social determinant of health inequalities.

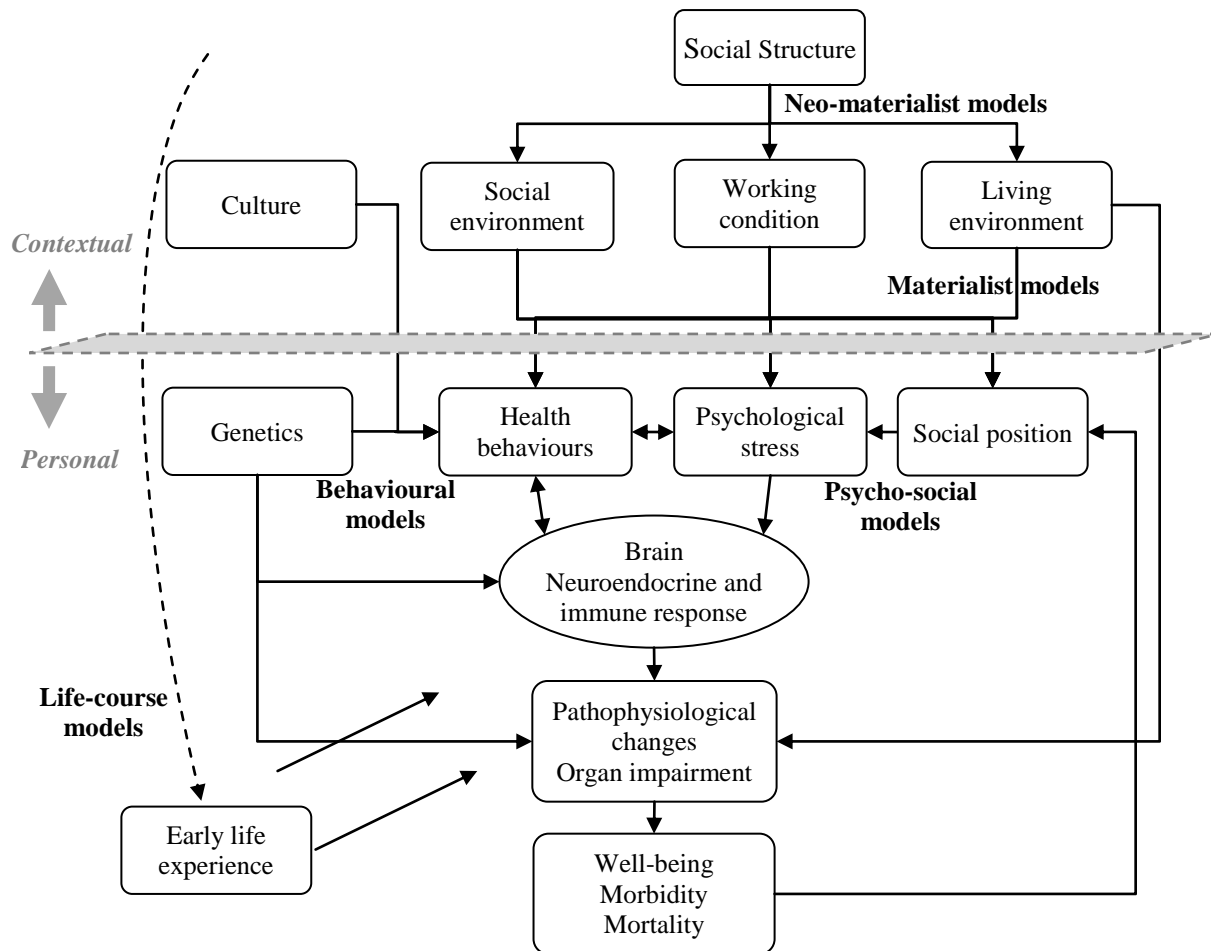


Figure 3.6: An integrated framework of social determinants on health (After Marmot & Wilkinson, 2006, p9)

### 3.2 Debates on composition and context

The numerous social determinants and causal pathways of health have created difficulties and generated debates in combining theoretical models to explain health inequalities. The debates are largely over the impacts of composition and context. While the compositional explanation argues that it is “who people are” that determines health, the contextual explanation on the other hand argues that it is “where people live” that determines health (Shaw et al., 2002). Table 3.3 illustrates various compositional and contextual factors and their respective assumptions.

Each of the four identified pathway models contains compositional and contextual components. Table 3.4 shows how individual and contextual factors may affect health inequalities through different casual pathways. The importance of distinguishing the impacts of individual or contextual factors is that predispositions for one or the other may lead to different policy implementations and health promotion strategies (Raphael, 2004).



	<i>Compositional (personal)</i>	<i>Contextual</i>
<b>Factor</b>	Individual characteristics, genes, life style choices, age, sex, race/ethnicity, health behaviours, socio-economic positions, income and education.	Employment opportunities, education provision, housing, transportation, retail, recreation facilities, land use, residential segregation, health services, health policies, environment hazards, social support, social networks, cultural norms, and climate.
<b>Assumption</b>	Population health is an aggregate of the health of individual members. There are no effects of the environment in which a person lives over and above their individual characteristics.	Social, economic, physical, and geographical environments have particular impacts on health inequalities despite different individual compositions.

Table 3.3: Compositional and contextual factors and influences

<i>Models</i>	<i>Levels of explanations for health inequalities</i>	
	<i>Compositional</i>	<i>Contextual</i>
Behavioural	Personal genetic composition, coping skill, life style choice and habits cause health difference.	Cultural norms affect life style choices, education, and behaviour. Unequal societies cause poor quality social relationships and make less well-off people feel less valued and take risky behaviours.
Materialist/ Neo-materialist	Low personal income influences home conditions, living environment and access to food and health care.	Risks, experiences, and opportunities are socially constructed. Lower level of public services, such as schools, libraries, public housing, transportation, and health care causes lower health status. Unequal social and political structures lead to unequal distribution of resources.
Psycho-social	Stress is a property of individuals. Low rewards at work induce stress, which affects the endocrine system, HPA activation, and immunity.	Increased social inequalities and lower-paid jobs lead to lower control of the less well-off at work, which makes people feel they have lower status and affects their immune and cardiovascular systems. Unequal income distribution and widened hierarchy reduce trust, weaken social cohesion, and damage quality of social relationships.
Life-course	The accumulation of personal risk factors, such as bad behaviours, stress, and material poverty, causes health inequalities.	Accumulated impacts of labour market, education system, child care, social support, social inequalities and living and working environments over the life span lead to health differences.

Table 3.4: Compositional and contextual explanations on health inequalities

When people believe that health outcomes are personal responsibilities due to individual personalities, coping skills, behavioural choices, and social achievements and status, they are correct to some degree. However, it is increasingly recognized by researchers and the general population that health outcomes are also community responsibilities. Communities are responsible when adults don't have safe places to walk, jog, or ride a bike, or children don't have accessible places to play; when school lunchrooms and work-place cafeterias do not provide healthy and appealing food choices; when teenagers in a neighbourhood smoke together and explore together other health threatening

behaviours; when people do not receive education on healthy behaviours and beneficial life styles; when there is no clean drinking water or a neighbourhood is exposed to pollutions; when neighbourhoods are threatened by crimes; and when local public libraries, transportation, and health care services are not easily accessed. There are many more community characteristics that can be identified as being responsible also for personal health.

Hence, compositional and contextual factors play key roles in determining health inequalities (Kawachi and Berkman, 2003). The question is what portion of these inequalities is accounted for by composition and what portion is attributable to context. Some studies suggest composition has a stronger influence than does context (Mitchell et al., 2000a; Mitchell et al. 2000b), while others suggest contradictory results (Macintyre and Ellaway, 1998; Emslie et al., 1999). These studies indicate that the balance between compositional and contextual factors is different among different places, groups, and health outcomes. To understand how contextual factors may influence health inequalities over space, attention is turned in the next section to theories that explain the geographical variations of social determinants and their potential impacts on health inequalities. By asking questions such as why people are healthier in some places than in other places? and how characteristics of place and space exert their impacts on people's health? It is possible to discover the underlying causes of health inequalities and thereby make health gains.

### **3.3 Geographical and planning perspectives on health inequalities**

Social and spatial factors are closely linked. Social activities shape places, and places in turn play important roles in constraining, distributing and regulating social activities, and mediating social impacts. To understand the role that place and space play in shaping health inequalities, studies in health geography and health planning have brought together theories from various disciplines to generate a systematic understanding of the spatial variations of health. Generally, three perspectives, namely the ecological, socio-economic, and psycho-social perspectives, can be identified to explain the geographic effects on health inequalities. These views, by focusing on the impact that place makes to the distribution of mortality and morbidity, may shed light on alternative health intervention policies and planning to tackle health inequality problems identified in the earlier discussion.

#### **3.3.1 Ecological landscapes**

The concept of an “ecological landscape” originated from human disease ecology and the effects of landscape epidemiology on infectious disease mapping. Research related to this concept is particularly concerned with the spatial patterning and diffusion of risk factors for diseases in the physical environment, and the biological and spatial attributes of human populations that influence

these risks (Barrett, 2000). It argues that spatial and group variations of health are due to varying exposures to risk factors, such as environmental pollution (air, water, noise, radiation, and chemical), climate change, risks of accidental injuries, or housing quality. This type of theory is consistent with the direct purely materialist model discussed earlier. It has been used to explain the biological and chemical pathways of risks and spatial patterning of particular diseases in human populations (Gatrell, 2002; Meade and Earickson, 2000). Geography as a spatial science contributes to knowledge in this area by describing and analysing the uneven spatial distribution of populations in relation to variable environmental risks and the consequent patterns of disease incidence.

The interplay of risk exposure and environmental impacts on health is through the triangulation of three interrelated groups of factors, namely population, habitat and behaviour (Figure 3.7) (Meade and Earickson, 2000). *Population* refers to both composition (such as age, sex, and genetic compositions), and processes (spatial distribution and patterning). *Habitat* includes both physical and social environments. *Behaviour* represents social organization, social norms, beliefs and technological resources of societies. The interaction of these three groups of factors determines health inequalities in space. The ecological perspective focuses mainly on the modeling of varying exposures of populations to physical environmental risks and the resulting spatial patterns of disease. The social aspects of disease risks are considered later in other geographical perspectives.

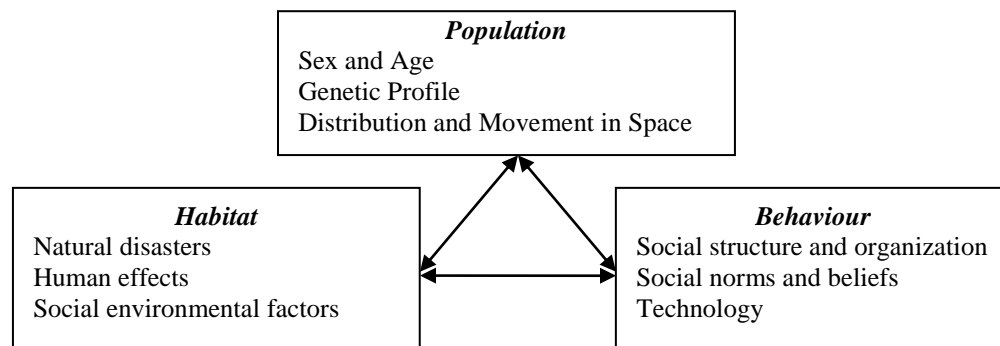


Figure 3.7: The framework of human disease ecology (After Meade and Earickson, 2000, p25)

The framework presented in Figure 3.7 is consistent with the earlier discussion that health variation is associated with both population composition and context. The population composition explanation in this setting argues that if people in an area are markedly more unhealthy compared with people in another area, this might simply be because the population in the worse off area contains a higher proportion of individuals who are at greater risk of ill health (due to genetic composition for example). The main role that environment plays in varying patterns of disease is in differentiating genetically related people (McGuffin et al., 2001). Many diseases, such as cancers, diabetes, and mental diseases, and behavioural traits, are to some extent genetically determined. What

the environment and geography do is to separate human populations with physical distance and topographical barriers and create distinct gene pools. The physical and social environments create “containers”, which constrain the population composition, direct population movement, and limit population interaction and the selection of marriage partners (Overall and Nichols, 2001). All of these factors affect the opportunities for genetic mixing and consequently for health variation.

In contrast to the genetic mixing perspective, studies of immigration have suggested that over time the epidemiological profile of immigrant populations becomes more similar to that of the host population and the second generation tends to have epidemiological characteristics closer to those of the host population than their parents (MacCredie et al., 1999; Balzi et al., 1995). This leads to the contextual ecological explanation that spatial variation of health is determined by people’s exposure to different physical environmental risks since neither the ethnical cultural influences nor the racial genetic composition may explain this phenomenon.

The spatial processes of population movement, such as internal and external migration, are important to exposure of population regarding environmental risks in particular places. Relocation of population within urban areas, influenced by individual motivation and affordability and socio-economic and spatial constraints, results in the spatial sorting or clustering of different segments of the population. This spatial clustering is particularly harmful for people living in deprived areas. The poor physical environment and potentially high pollution typically formed in these areas, combined with the relatively low social status of local residents, create double jeopardy to damage their health.

The spatial movement of population is also crucial for the diffusion of some infectious diseases, such as HIV. The geographical proximity of infected and non-infected populations can be powerful predictors of the chances of infection (Meade and Earickson, 2000; Cliff and Haggett, 1988). Spatial modelling of disease diffusion is especially relevant to this kind of study (Cromley and McLafferty, 2002). It argues that diseases spread outward from their sources in ways that can be predicted from the spatial distribution and movement of populations. People who are closer in space are more likely to communicate than those who are separated by large distances. Geographical barriers, such as rivers or mountains, may also inhibit the communication of population and diseases. Therefore, disease diffusion can possibly be described by the two factors of distance and barriers.

Besides geographical barriers, the hierarchical order of city systems also creates social barriers and makes disease diffusion more complex (Meade and Earickson, 2000). Rates of communication are more rapid between large cities than between smaller towns. It is very likely that new infections are introduced first to cities with the largest populations and the greatest rates of contact with outside regions, and then spread out to peripheral smaller towns. Based on population distributions and likely migration or communication patterns, it becomes possible to summarize the likelihood of diffusion of

infectious diseases mathematically, providing evidence for disease prediction and prevention.

In addition to population factors, the uneven distribution of environmental risk factors also influences variations in potential risk exposure. For infectious diseases, the diffusion follows Agent-Vector-Environment-Host pathways (May, 1958). A range of natural environmental factors, such as climate, topography, vegetation and stagnant water, may influence the transmission pathways. Human factors, such as behaviours in protecting from exposure to vectors and the human composition and movement discussed above, influence the susceptibility and spread of diseases as well. All these lead to the spatial variation of potential exposures to infection risks.

For non-communicable diseases, environmental impacts and geographical processes are more complex than those of infectious diseases because of their chronic nature and less clear pathways. Studies addressing these issues are often called environmental epidemiology, defined as “the study of the effect on human health of physical, biological and chemical factors in the external environment” (Needleman, 1997, p263). The purpose of these studies is to discover the health risks associated with aspects of the physical and biological environment. A range of possible health hazards in the environment is associated with health variations, such as the association between air pollution and respiratory diseases. However, the exact physiological processes by which a certain environmental hazard may affect human health are not always well understood. This is due to the complexity of potential environmental risks, which may be broad and operate at many scales. Thus, the associations found between health and environmental factors often provide only limited understanding about why these associations occur. This situation is referred to as “black box epidemiology” (Figure 3.8) (Curtis, 2004). As such, it is part of the goal of this thesis to discover what factors and relationships exist inside the black box. The methodological discussions on identifying the causal pathways are presented in Chapter 4.



Figure 3.8: Black box epidemiology (After Curtis, 2004, p175)

The uneven distribution of environmental risks also raise a further major complication, namely environmental injustice (Jerrett et al., 2001). This is evoked by the fact that areas with high environmental risks are often occupied by socially disadvantaged populations. For example, industrial sources of pollution are often located near socially and economically disadvantaged communities, such as aboriginal reservations. This systematic environmental injustice faced by disadvantaged social groups can be explained by the uneven distribution of power between social groups and the consequent weak environmental protection policies. This issue is discussed later in the next section.

A conceptual framework of the ecological perspective for unequal environmental risks is summarized in Figure 3.9 (Curtis, 2004). In this framework, the key factors of population, habits and behaviour demonstrated in Figure 3.7 in conjunction with environmental justice lead to socially and spatially unequal exposures to physical environmental risks of health. The interplay of these varying exposures, together with other social and personal risks, generates unequal health outcomes for different population groups in different areas. This framework provides the theoretical basis for further exploration of the pathways from environmental risk factors to health inequalities.

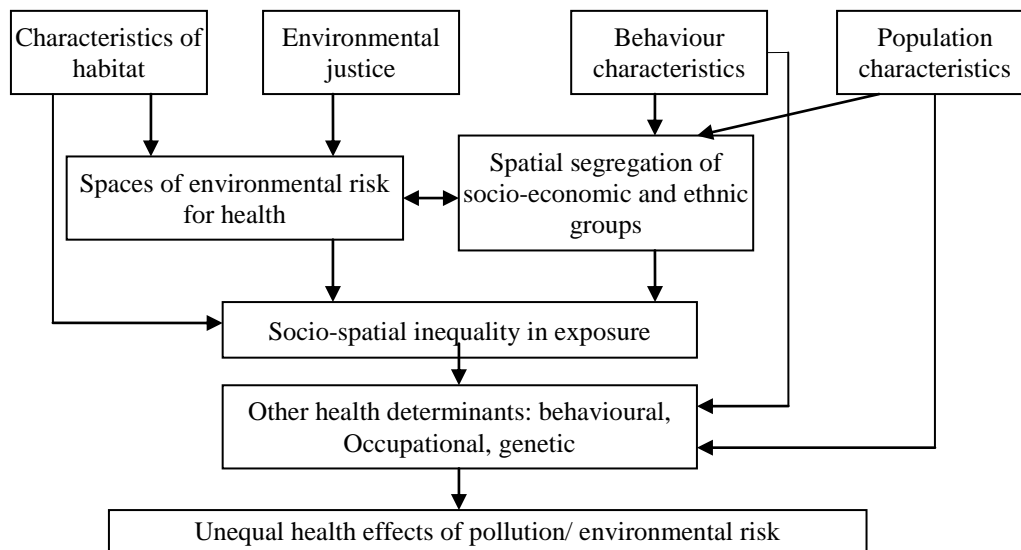


Figure 3.9: Conceptual framework for socially unequal environmental risk (Adapted from Curtis, 2004, p188)

### 3.3.2 Theories about space and place in social relationships

While the ecological landscape mainly focuses on the links between the physical environment and the health risks of populations, it lacks power in explaining how the unequal distribution of population and environmental risk factors are manipulated by various social and economic forces. Some theories derived from social science, on the other hand, emphasize the interplay between social relations and space in the production of health inequalities. The focus of these theories is on how social interactions and health-related services, resources, and products are constrained, constituted, and mediated through space to generate health variations. Two perspectives can be particularly useful in explaining these social and spatial relations, namely theories of structuration and theories of political economy.

### 3.3.2.i Theories of structuration and the geography of power

Theories of structuration emphasize the interaction between individuals and their socio-geographical context. The interactions of individuals shape, and are shaped by, social structures, which comprise principles of social organization, rules and systems in space and time (Giddens, 1984). A person's daily actions are taken at a certain time and in a certain space, which bring the person into routine contact and interaction with other people in a variety of hierarchical "locales" such as homes, schools, hospitals, work places, neighbourhoods and cities. A person's social practices interact with social structures of locales in ways that are not always obvious to them. They follow the rules and norms of the locales consciously and subconsciously to perform their routine social practices, thus helping to maintain and shape social structures. The established social structures, in turn, frame the actions of individuals.

Social structures may be shaped and extended through the coordination of natural and human resources over time and space. As different social groups struggle to obtain control and subdivide resources, social structures are distinguished and resources are ordered along social lines based on the power of each group. The ordering of resources in time and space puts constraints on individual actions, which helps to maintain the power relations between dominant and subordinate groups.

Socio-geographical spaces operate as "power containers" in this mechanism. Socio-economic or political power is exercised through the processes of territoriality and surveillance over spaces. Territoriality is established when area boundaries are plotted out to control access for others (Sack, 1986). Surveillance is exercised by setting up rules and supervising people's performance to maintain them (Dandeker, 1990). Through territoriality and surveillance, society becomes a "disciplinary cage" (Foucault, 1979), which offers safety and security, but also reinforces social structures and social divisions. Thus, social and spatial inequalities are maintained and reproduced through the struggle for power over different sites between more and less powerful groups in society. The result of these struggles reinforces a sense of difference between "the same" and "the other", and creates social and spatial distance between them. As individuals of the same groups participate in shared time-space paths, "acceptable norms" are established (Butler and Bowlby, 1997), social and spatial boundaries are set up, and group differences in power and access to resources are perpetuated (Dear et al., 1997). Individuals are therefore deemed to be "the other" when they are not fully involved and cannot conform to the acceptable norms of a given group.

Different groups, by occupying different territories over space, and accessing different resources, have different degrees of exposure to health risks. This is especially true for groups that are frequently subordinated in society, including women, children, the elderly, the homeless, minority ethnic groups, low-income or unemployment groups, people with chronic illness or disabling

conditions, or those who are marginalized because of their behaviours, such as heavy smokers, alcoholics, and illicit drug users. These groups have less control over resources than others and are typically disempowered in decision making processes. They are often disadvantaged by their material poverty, and marginalized or isolated in space and place. The material deprivation and spatial segregation produced by the processes of social and spatial exclusion processes create large health disadvantages for these groups.

An example of detrimental effects for local residents is the urban desertification process in the inner city of New York between 1970 and 1980 (Wallace and Wallace, 1997). Desertification took place in areas occupied by large concentrations of poor populations. Due to their inner city locations, these areas usually have potentially high commercial value or may impede the development of their surroundings. The struggle for control over these areas between local residents and other politically powerful groups had the result that some of the local communities were forced to move or were displaced to other locations. During the desertification process, basic services, such as fire fighting and policing, were literally withdrawn from these areas, resulting in a steady decline of the social and physical environments. Whoever could afford to move vacated these areas, leaving a growing concentration of the most disadvantaged and excluded groups. The remaining population experienced a steadily worsening health situation, on top of the high risk of illness and lack of access to health services they had already experienced in their lives. This process, imposed by more powerful social groups, created socio-spatial exclusion and marginalization of those in the weakest social position, and consequently damaged the health of community members (Andersen, 2003).

Thus, hegemonic power structures exert detrimental effects on the living conditions and the health of disempowered populations through the processes of social and spatial polarization. Based on these relationships, the processes of structuration, exclusion and disempowerment provide plausible explanations for health inequalities.

### 3.3.2.ii Theories of political economy and the landscape of collective consumption

While structuration theories explain how social and spatial “structures” are formed, changed or stabilized by struggles between different power groups over resources, the political economy perspective attempts to explain why and how socio-economic “classes” or different “power groups” are formed, maintained, and reproduced, and how inequalities in social status, income, and service provision operate in social and spatial structures to produce health inequalities.

The term “political economy” implies that politics and economies are integrally related. Society is treated as a whole, which “is shaped, but not determined, by the mode of production, that is, by the means of producing and reproducing for socially determined needs” (Armstrong, et al., 2001, p.VII).



Hence, societies respond to the needs of their members in order to ensure human survival and reproduction, and put mechanisms in place in order to sustain human life. The dominant mode of production in industrialized countries is capitalism, under which production is organized through the buying and selling of land (space) and labour (people) to create more capital. The entire dynamic of the capitalist mode of production is geared to continue the accumulation of capital (Pacione, 1987). Thus, society is divided into major classes of labourers and owners, who are in continuous conflict because capital accumulation can only be achieved through the exploitation of labour.

In the last several decades, industrialized countries have been in the process of transformation from modernism to post-modernism, from Fordist to post-Fordist modes of production, from organized to disorganized capitalism and from localized economies to globalization. The globalization of the world's economy, or neo-liberalism, emphasizes the market as a source of wealth and justice, and assumes that a "free-enterprise" economy will produce the greatest good for all (Coburn, 2000). This view suggests that neo-liberalist approaches to economic policies will result in the accelerated accumulation of capital. However, this also creates a widening in the gap between socio-economic classes and induces social and spatial polarization through rearrangement of the labour force and restructuring of social classes (Martens and Rotmans, 2002). The "smooth" socio-economic gradients in the past are now changing to two sorts of conditions, namely affluence and poverty on the one hand, and participation and exclusion on the other hand (Byrne, 1997). These increased income and social inequalities create potential conflicts between socio-economic classes and are associated with increasing health inequalities. This phenomenon is already explained by both the psycho-social model and the materialist/neo-materialist model discussed earlier.

These social divisions are coupled with spatial divisions. The capitalist mode of production generates uneven development and inequalities of wealth between geographical areas. The spatial patterns may take various forms based on the organization of production (Curtis, 2004). In the classic Fordist model, industry is designed for mass production, concentrated in large factory settings and usually located within urban cores. Different types of work are spatially divided and standardized. This enables the profitable production of standardized goods at relatively low prices. Post-Fordist modes of production, transformed from mass production to new technologies and service industries, are more flexible and specialized. Industries tend to have multiple production units which are sparsely distributed in space, with associated industries located close together, often in accessible but less urbanized locations.

The spatial organization of both Fordist and post-Fordist forms of production is associated with the spatial division of wealth and the spatial separation of different classes in residence, often referred to as socio-economic segregation (Pacione, 1987). This segregation, on the one hand, facilitates the

reproduction of social classes for maintaining an appropriate labour power structure and, on the other hand, according to the structuration theories discussed earlier, it forms the social milieu and community organization that resists threatening intrusions and promotes avoidance of undesired contact. However, such segregation is detrimental to the poor in terms of both quality of life and health due to the associated poor living conditions, low quality education, fewer job opportunities, and less influential social contacts. The shift from Fordist to post-Fordist modes of production also results in a decline in “old industrial” areas. These areas were once successful, but are now disadvantaged, combined with the marginalization of the local residents. The detrimental health effects of living in deprived areas highlight the need to address health inequalities by considering not only individual poverty and wealth, but also the socio-economic and spatial structures in which people live and work. These structures cannot be effectively influenced by individual action alone. They should also be the responsibility of local and national governments.

While macroeconomic processes create conflicts between the production of goods and the improvement of well-being, they tend to be internationally organized. National governments play less and less of a role in direct intervention in these processes. Hence, the widening gaps between the rich and the poor are not likely to be reversed without dramatic structural changes. However, governments do play important roles in moderating the negative effects of widening social gaps. Different governments operate different “regimes” (Stone, 1993), or modes of regulation, in order to achieve economic stability and the reproduction of the capitalist political economy. They operate through economic and social policies and regulations, through electoral systems, through welfare benefits, and through the provision of a wide variety of services (such as education, housing, transport, and importantly health care). All of these roles influence the operation of markets, maintain the relationships between socio-economic classes, and keep a balance in society.

In this sense, the provision of state health care is seen as one way in which societies maintain and regulate capitalist political economies. As defined by the World Health Organization (1998), health is treated as a collective good and a basis for social and economic activities. Collectively, health services ensure “fitness to work” and help to reproduce the labour force. They also help to compensate workers for their relatively weak position in the political economy and therefore maintain the political stability of the system. With the widening gaps between socio-economic classes in the globalized economy, the provision of public health services and other collective goods has become increasingly used to mitigate the increasing impact on health and welfare of capitalist modes of production, and maintain capitalist societies as viable political economies (Sanger and Sinclair, 2004). However, various social, political, and geographical reasons determine that the provision and consumption of health care are not equally distributed among individuals and geographical areas,

which directly influences health inequalities.

On the provision side, health service distribution is influenced by the political and administrative structures that are in place. In the context of a city, for example, the decisions that determine the direction of local policy and the deployment of resources to different areas and population groups are constrained by the organizational structures of local governments. Decisions are made partly with the aim of preserving the existing organization, rather than necessarily reflecting public “need”. This may be the outcome of political struggles between different classes in society and a reflection of the dominance of the capitalist socio-economic structure. These decisions result in the local variation of health infrastructures including their availability, accessibility, accommodation, affordability and acceptability (Penchansky and Thomas, 1981). Taken together these factors eventually affect a person’s access to health care.

On the consumption side, the pattern of consumption is not merely determined by personal needs or wants. Needs are socially structured and result from the value system of society and the capitalist production mechanism. Similar consumption patterns may therefore be shaped within similar social groups or social classes. In a geographical context, based on the structuration theory discussed earlier, different social groups occupy different locales which shape the lifestyles and social norms of the local people. The social structures of these locales harmonize consumption practices of the “same” and differentiate these shared patterns from “others”. The consumption of health-related goods and services, including informal and voluntary health care, is likely to depend on the social composition of the local community, and local cultures of health-related behaviours and beliefs. Thus, different consumption beliefs and behaviours separated by locales or socio-geographical spaces may end up causing spatial variation in health care access as well.

When different groups struggle to impose their preferred consumption over a particular space, conflicts may emerge in a way that favours the groups who possess power (Armstrong et al., 2001). The consumption pattern of socially advanced groups is usually manifested in ways that exclude or disadvantage less influential groups. As urban space is first occupied by facilities offering exclusive consumption for the advantaged groups, less space is left for other forms of consumption that are important for less powerful groups, such as local grocery stores, street markets, public open spaces, free-access play spaces, and public transit. Through territory and surveillance, the consumption of high cost and high quality facilities has constrained access for the more powerful groups, which limits their ability to benefit less powerful groups from these aspects of lifestyle, and to improve their health.

Since physical infrastructures are fixed in space, the spatial distribution of health facilities also influences health care use. In organizing the spatial distribution of health services, central place

theory invokes the idea of a hierarchy of service centres, such that higher order centres provide a larger range of services over a wider area, while lower order centres offer services to a local catchment (Carter, 1981). The gravity model extends this idea by replacing the fixed area of catchments with a distance decay effect in the use of a service facility (Congdon, 2000). Since the rates of hospital visits tend to decline as distances from the hospital increase, the gravity model estimates utilization as inversely proportional to the squared value of distance between users and facilities. Other factors, such as physical accessibility, due to lack of transport or infirmity, also limit the accessibility to health services, especially for remote rural area residents and individuals with restricted physical mobility. Geographical distance and barriers do matter in health service uses. However, spatial accessibility is often not the dominant factor influencing the distribution of health facilities.

Drawing linkages between the broad social, political, economic, and geographical processes and health variations, this section has discussed how social structures, political organizations and economic modes of production result in socio-economic inequalities, inequalities in the distribution of resources, and varying patterns of consumptions over space. The processes through which health determinants go down the social hierarchy from broad structures to local geographical areas are explained by structuration theory and political economy perspectives. However, distinctive neighbourhood characteristics that are shaped, but not determined, by social structures, such as social capital, are ignored in these explanations. Social capital exists within each neighbourhood independent of the purposes or reasons of the original gathering of local residents. It directly influences the health and quality of life of local residents and therefore needs to be further addressed.

### **3.3.3 Social capital and community**

Social characteristics vary systematically across local communities or neighbourhoods. Although causality and magnitude are still at issue, community characteristics, such as concentrated poverty, segregation, family disruption, residential instability, and poor quality housing, are closely linked to health problems. While structuration theory and the theories of political economy explain the systematic variation of these characteristics across communities, they are not sufficient in uncovering the underlying causes or mediating mechanisms that adhere to communities, such as social cohesion, social networks, and informal social controls. One explanation to address such social processes is the theory of social capital.

Social capital is defined as the “features of social organization, such as the extent of interpersonal trust between citizens, norms of reciprocity, and density of civic associations, that facilitate cooperation for mutual benefit” (Putnam, 1993, p36). Social capital is generated through reciprocal

obligations among neighbourhoods, information exchange, intergenerational closure, and voluntary associations (or social networks). It encourages cooperation and facilitates social cohesion, which refers to the absence of social conflict coupled with the presence of strong social bonds and mutual trust (Kawachi and Berkman, 2000).

Social capital or trust is positively linked to health (Kawachi et al., 1999). However, local social ties or local social networks do not necessarily translate into high social capital at the neighbourhood level. Disadvantaged urban neighbourhoods with strong social ties may actually impede social organization (Wilson, 1996). This high degree of social integration within neighbourhoods isolates the disadvantaged groups from contacts in broader mainstream society and makes them have low levels of informal social control over their immediate environment.

To address the seemingly paradoxical relation of social cohesion and social ties, the concept of collective efficacy has been proposed to complement social capital (Sampson et al., 1997). Whereas social capital refers to the resources inherent in social networks, collective efficacy refers to a task-specific construct that facilitates social control without requiring strong ties or associations. It emphasizes the shared belief in a neighbourhood's collective capability for action to exercise control over a neighbourhood issue. It therefore leads to an active sense of purpose and engagement of local residents to intervene in support of public order in their neighbourhood. Thus, both social control and social cohesion can be achieved with high levels of collective efficacy, which is positively associated with lower rates of violence and delinquency, less material deprivation of neighbourhoods, and high levels of social integration and residential stability.

Both social capital and collective efficacy are closely linked to population health. People are likely to be happier and healthier when they live in neighbourhoods where there is a sense of pride and belonging. In contrast, people who live in neighbourhoods with low levels of social capital and collective efficacy are likely to be depressed and have low control over their environment, and consequently have poor health. Thus, social capital or collective efficacy can be seen as a remedial factor likely to reduce the negative psycho-social effects caused by social polarization and work stresses. Social capital thus provides an alternative mechanism for relief of social and spatial health inequalities.

Overall, while the causal models discussed in Section 3.1 focus on the “pathways” of health inequalities, the geographical perspectives examined in this section emphasize the “interpretation” of how social determinants exert their impacts on health through space. These ecological and socio-geographical theories, while useful in an explanatory sense, are narrative in nature and may not be testable. The challenge, therefore, is to establish how to translate these theories into testable models and how to find a place in the pathway models to integrate effectively the spatial impacts of health, so

that the social and spatial pathways of health determinants can be effectively discovered. The next section, bearing this in mind, based on research discussed in Chapters 2 and 3 and the availability of data collected, proposes hypotheses of social and spatial determinants of adverse birth outcomes that are subsequently tested for this thesis.

### 3.4 Research hypotheses

As stated in Chapter 1, the goal of this thesis is to clarify the role that place and environment play in the causal pathways of social inequalities that lead to adverse birth outcomes, specifically low birth weight and preterm births, in a developed society. At the same time, the research seeks to identify possible area-based and community-oriented interventions or policies to reduce these inequalities and reduce adverse birth outcomes. Based on the above discussion and a first pass assessment, a number of hypotheses are proposed, according to the local situation.

#### *Hypothesis 1: Personal vs. Spatial*

The general pathway theories and the geographical perspectives of health inequalities discussed above and the evidence provided in Chapter 2 suggest that space and place may play an important role in shaping inequalities of adverse birth outcomes. Thus, the first hypothesis to be investigated is framed as follows:

The Null Hypothesis,  $H1_0$ , is: *No significant spatial differences of adverse birth outcomes exist among different neighbourhoods. In other words, no adverse birth outcomes are clustered over space.*

The Alternative Hypothesis,  $H1_1$ , is: *Significant spatial differences of adverse birth outcomes exist among different neighbourhoods. In other words, adverse birth outcomes are clustered over space in certain neighbourhood areas.*

The alternative hypothesis may be manifested in two ways. First, considerable variations of adverse birth outcomes may exist among neighbourhoods over the study region compared to variation of adverse birth outcomes for individual mothers. Second, spatial autocorrelation may exist among neighbouring areal units. In other words, areas of significantly high concentrations or “hot spots” of adverse birth outcomes may exist within the study region. The confirmation of this hypothesis provides the basis for further neighbourhood-level investigation.

#### *Hypothesis 2: Compositional vs. Contextual*

The second hypothesis concerns the need to distinguish the contextual effects of neighbourhoods from their compositional effects. In other words, the question to be examined is whether the spatial difference or neighbourhood difference of adverse birth outcomes expressed in  $H1_1$  is a result of the composition of personal risk factors in each neighbourhood, or a result of neighbourhood

characteristics. Since the compositional effects at the personal-level are already confirmed by other research and, given the existence of spatial variation among neighbourhoods, the second hypothesis is:

The Null Hypothesis,  $H2_0$ , is: *There are no direct contextual impacts of neighbourhoods on adverse birth outcomes. Spatial variations are due solely to the compositional effects of individual characteristics of mothers.*

The Alternative Hypothesis,  $H2_1$ , is: *Contextual risks (neighbourhood characteristics) contribute to spatial variations of adverse birth outcomes after controlling for personal compositional risks.*

The alternative hypothesis,  $H2_1$ , arises because the confirmation or rejection of this hypothesis will have direct impact on the focus of policy implementation and health promotion strategies. A confirmed association with neighbourhood, including socio-economic and ethnic composition, living conditions, environmental pollution, health services, basic facility provisions and social capital, may provide evidence supporting direct interventions in these risks at the neighbourhood level. However, this does not necessarily mean that neighbourhood risks diminished by controlling for personal risks can be ignored in community health interventions. These neighbourhood risks may affect personal-level risks through concentrating personal risks or through modifying the effects of personal risks. Understanding these mechanisms will help identify intervention targets at both levels. This consideration leads to the third and fourth hypotheses that examine the pathways of adverse birth outcomes.

### *Hypothesis 3: Psycho-social vs. Material*

In addition to known direct medical risks to the foetus, neighbourhood-level risks also affect adverse birth outcomes through maternal medical, psycho-social, and behavioural risks. Even though some of the neighbourhood-level risks may be completely mediated by personal-level risks, it is important to understand how the mediation effects operate and what pathways they follow. This will provide evidence supporting the direction and focus of interventions to reduce adverse birth outcomes.

As discussed in Chapter 2, neighbourhood-level socio-economic risks have been consistently associated with adverse birth outcome differences. However, the specific mechanism or mechanisms that are responsible for variations between neighbourhoods of different social classes remain unknown. Psycho-social explanations suggest that the differences are caused by associated psycho-social factors or stressors, whereas a materialist explanation emphasizes that birth outcome differences are the results of material differences or different purchasing or controlling powers between social classes.

It can be observed from Figures 3.3, 3.4, and 3.6 that the processes of socio-economic impacts on health are complex, involving both upstream and downstream influences. The focus in this thesis

is the downstream influence of neighbourhood socio-economic effects on adverse birth outcomes. Given the association between neighbourhood socio-economic status and adverse birth outcomes, the hypotheses to be tested are framed as:

The Null Hypothesis,  $H3_0$ , is: *Neighbourhood social-economic status does not affect mothers' personal-level risks for low birth weight and preterm births. Thus, the risks of neighbourhood social-economic status on adverse birth outcomes are not mediated by any personal-level risks.*

The First Alternative Hypothesis,  $H3_1$ , is: *Neighbourhood socio-economic risks affect adverse birth outcomes through the relative social position of mothers among neighbourhoods. The effects of neighbourhood socio-economic status on adverse birth outcomes are mediated by maternal psycho-social or behavioural risks due to different social positions defined by neighbourhood socio-economic status.*

The Second Alternative Hypothesis,  $H3_2$ , is: *Neighbourhood socio-economic risks affect adverse birth outcomes through associated material conditions. The effects of neighbourhood socio-economic status on adverse birth outcomes are mediated by maternal personal socio-economic risks, health risks, and living conditions.*

The first alternative hypothesis,  $H3_1$ , favours the psycho-social model in that a neighbourhood with low socio-economic status may put psycho-social stresses on mothers living there and may consequently lead to adverse health behaviours during pregnancy. This hypothesis also favours the geographic perspective of social capital and community. Based on this hypothesis, strong neighbourhood social capital should have a buffering effect on adverse birth outcomes. On the other hand, the second alternative hypothesis,  $H3_2$ , favours an indirect materialist model which suggests that low neighbourhood socio-economic status leads to low individual socio-economic status, poor living conditions, and poor personal health conditions that eventually lead to adverse birth outcomes. This hypothesis also favours the pure direct materialist beliefs (as stated in the Alternative of Hypothesis 2,  $H2_1$ ) and the geographical perspective of ecological landscapes, in that neighbourhood conditions, such as housing, health services, and food provisions, have direct impacts on adverse birth outcomes. These two alternative hypotheses do not necessarily exclude one another. The interrelations of material and psycho-social factors make it highly possible that both of them are simultaneously at work.

#### *Hypothesis 4: Personal vs. Cultural*

In addition to neighbourhood socio-economic status, different neighbourhood ethnic groups also exhibit varying degrees of associations with adverse birth outcomes. The behavioural models discussed in Section 3.1.1 suggest that these associations may be due to behavioural differences. However, there is also debate over the roles of personality (genetic composition) and culture. The



behavioural models suggest that behavioural risks can be either the result of group influences (norms, beliefs, and life styles among ethnic and other social groups), or the consequences of personality. Given the established association between neighbourhood ethnic composition and adverse birth outcomes, the fourth hypotheses to be tested are:

The Null Hypothesis,  $H4_0$ , is: *There is no direct or indirect impact of neighbourhood ethnic composition on adverse birth outcomes or on related health behaviours during pregnancy.*

The First Alternative Hypothesis,  $H4_1$ , is: *Neighbourhood ethnic composition affects mothers' adverse birth outcomes by changing mothers' health behaviour during pregnancy through cultural influences within their neighbourhoods.*

The Second Alternative Hypothesis,  $H4_2$ , is: *Neighbourhood ethnic composition affects mothers' adverse birth outcomes due to the inherited predisposition of specific group members to adverse birth outcomes or to adverse health behaviours during pregnancy.*

Since health inequalities among ethnic groups are often linked to socio-economic disadvantages, the currently established association between neighbourhood ethnic composition and adverse birth outcomes may be primarily due to confounding characteristics that are associated with ethnic status, such as socio-economic status. Through controlling neighbourhood-level socio-economic status, the null hypothesis,  $H4_0$ , can be tested by assessing whether the association still exists after controlling for the confounding influence of socio-economic status.

The first alternative hypothesis,  $H4_1$ , is based on a cultural explanation of the behavioural model. It emphasizes the important role that group culture plays in affecting people's beliefs and consequent choice of life-style choices and health behaviours. It also favours geographic perspectives about place and space relative to social position as described by the theories of structuration and the geography of power.

On the other hand, the second alternative hypothesis,  $H4_2$ , favours a direct behavioural explanation which assumes that certain groups of people are inherently predisposed to certain types of personal characteristics, such as the lack of coping skills to exert control over their circumstances, and because of these, they are more likely to have adverse health behaviours. The population composition explanation of the ecological landscape perspective discussed earlier argues that the environment and geography sort human populations according to physical and social distance, and the resulting barriers to interaction create distinct gene pools in different neighbourhoods. In addition to affecting maternal health behaviour during pregnancy, this hypothesis also favours a direct genetic impact on adverse birth outcomes.

It is usually difficult to test directly the association between genetic composition and health behaviours. However,  $H4_1$  and  $H4_2$  can be tested indirectly by controlling group or neighbourhood

level ethnic composition (controlling the cultural influence). If personal racial/ethnic status still affects behavioural influences on adverse birth outcomes, the possibility that inherited predisposition may play a role in affecting health behaviours, and consequently affect adverse birth outcomes, cannot be rejected. Otherwise, if personal racial/ethnic status does not affect behavioural risks or the outcome of births, it will be only the cultural influence at work.

#### *Hypothesis 5: Global vs. Local*

The above hypotheses are global in nature, seeking universally applicable associations between exposures to neighbourhood risks and responses in terms of adverse birth outcomes on the part of local women. However, due to the effects of local spatial autocorrelation or residential segregation, some of the exposures may only be harmful in certain neighbourhoods, especially those that are socially isolated and economically deprived. In this case, health promotion policies that are universally applicable to all neighbourhoods may not be effective. The impacts of personal risks may also vary from neighbourhood to neighbourhood due to different local situations or unidentified environmental risks. Based on this concern, the final research hypothesis to be tested can be stated as:

The Null Hypothesis,  $H5_0$ , is: *Risks of adverse birth outcomes, at both the personal and neighbourhood levels, have only universal or global impacts and there are no observable local impacts of neighbourhood characteristics on adverse birth outcomes.*

The Alternative Hypothesis,  $H5_1$ , is: *Personal and neighbourhood level risks affect adverse birth outcomes differently from neighbourhood to neighbourhood due to the complexity of local social contents or to the influence of unidentified local risks. That is, local “hot spots” of adverse birth outcomes exist due to the impacts of locally varying factors.*

This hypothesis suggests that it is necessary to use local spatial tests to identify the existence of clustering of adverse birth outcomes after controlling for identified risks, and the varying local influences of personal and community-level risk factors. If the existence of local spatial autocorrelation is confirmed, the results may provide information not only about where to act, but also about which actions should be taken and how to implement them for community-oriented health promotion programs to reduce adverse birth outcomes. This may help to overcome some of the reasons why the multi-component intervention programs discussed in Chapter 2 have failed. The identified local high incidence of adverse birth outcomes and local variation of personal risks may also provide shared targets to facilitate social control and social cohesion through the collective actions of local residents and multiple stakeholders in community health intervention programs. Through the engagement and participation of local high risk groups and multiple stakeholders, a bottom up approach may be initiated to achieve the goal of improving community health in general and reducing adverse birth outcomes in particular.

### **3.5 Summary**

In this chapter, four general pathway theories, namely the behavioural model, the psycho-social model, the materialist/neo-materialist model, and the life-course approach were introduced. Major geographical perspectives on health inequalities, including ecological landscapes, theories of space and place in social action, and the concepts of social capital were also discussed. Based on this general theoretical discussion, the research-based evidence provided in Chapter 2, and data collected for this thesis research, five testable hypotheses of neighbourhood impacts on adverse birth outcomes were derived. The purpose of testing these hypotheses is not only to prove general assumptions on the pathway approaches and spatial influences of adverse birth outcomes, but also to provide evidence for community health interventions. In next chapter, a quantitative analysis framework is established based on these hypotheses and the available data, available statistical methods are discussed, and suitable statistical models are established.

## **Chapter 4**

### **Methodology**

The objectives and hypotheses stated earlier in the thesis require the examination of geographical variations of adverse birth outcome incidence and variations in the associations between incidence and explanatory variables, both individual and contextual, to be established. These relationships are both global and local so that strategic policy making and health resources allocations can be made accordingly. Statistical methods, including multilevel spatial modeling, Bayesian spatial hierarchical modeling, disease mapping, and clustering analysis, are needed to test the hypotheses stated in Chapter 3. Since the proposed research uses secondary data obtained from various sources, preliminary geoprocessing analyses was required for data management, geocoding, spatial interpolation, and factor analysis. This chapter introduces currently available statistical methods for related analyses and discusses the advantages and limitations of each method, so that a quantitative analysis framework can be established and effective statistical methods can be identified.

#### **4.1 Statistical analysis framework**

This section presents a framework of the analysis procedures and statistical methods required to conduct the proposed analyses (Figure 4.1). Within this framework, several data handling methods were first utilized to obtain clean, geo-referenced, and consistent data for both personal-level incidence and potential personal and neighbourhood-level covariates of adverse birth outcomes (refer to Chapter 5 for details of data used in this thesis). Geocoding by street address and by postal codes was subsequently used to obtain the spatial locations of live births in the study areas. Spatial interpolation methods, such as Kriging, were then used to derive neighbourhood-level indicators from secondary data sources.

After the necessary data were prepared, global and local spatial analyses were conducted to test the five hypotheses stated in the last chapter. The first hypothesis was tested and potential spatial impacts on adverse birth outcomes were examined using various spatial statistical analysis methods, including an analysis of intraclass correlation (ICC), Bayesian spatial random effect models, and global spatial clustering analysis. These methods examined globally, from different angles, the existence and extent of spatial or neighbourhood impacts on adverse birth outcomes.

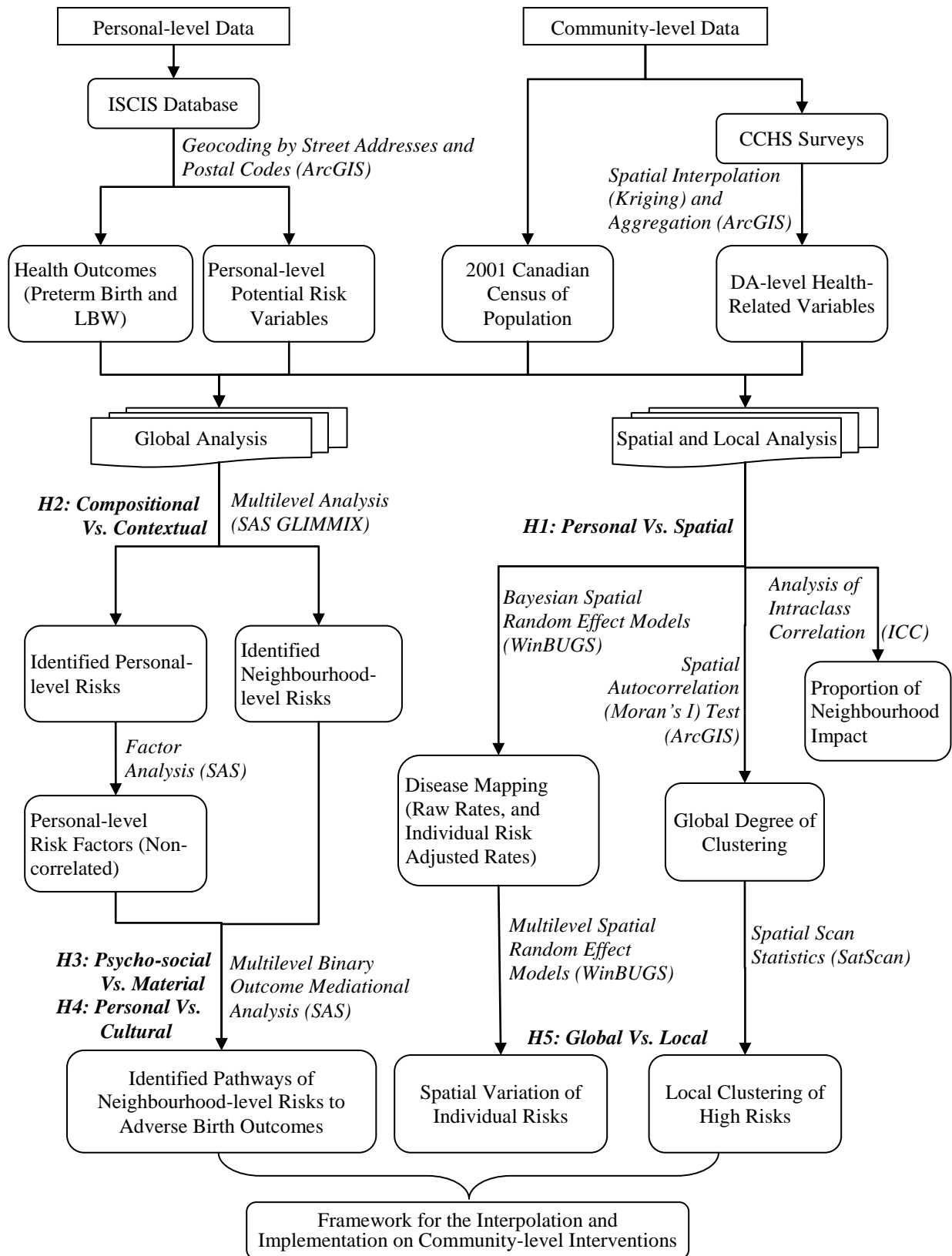


Figure 4.1: The quantitative procedure for adverse birth outcome analysis

Conformation of Hypothesis 1 and the existence of spatial variation of adverse birth outcomes among neighbourhoods allowed the examination, using multilevel modeling, whether the spatial impacts expressed in Hypothesis 1 were a result of compositional effects or contextual influences. Consequently, potential personal and neighbourhood level risks of adverse birth outcomes were also identified. To analyze effectively the pathways and local spatial impacts on personal-level risks and consequent adverse birth outcomes, a factor analysis was then conducted on the identified personal-level risks in order to extract personal-level common risk factors.

Multilevel binary outcome mediational analyses based on the extracted common risk factors and identified neighbourhood-level risks were then conducted in order to test Hypotheses 3 and 4. The results of these analyses provide evidence on a pathway or pathways through which neighbourhood risks may take to influence adverse birth outcomes. Given the existence of global spatial autocorrelation of adverse birth outcomes in the study region, local spatial analyses were also conducted using multilevel spatial random effect models and spatial scan statistics. Hypothesis 5 was then investigated and tested based on the resulting spatial clusters of adverse birth outcome incidence and spatial variations of the effects of personal risks on adverse birth outcomes.

Based on the analytic methods described in Figure 4.1, this chapter is organized into three sections, namely global analysis, local spatial analysis, and preliminary analysis. Within each section, the choice of each statistical method is explained by comparison with other available methods and details of the chosen methods are introduced. Specific statistical models for hypothesis testing are also constructed following the discussion.

## **4.2 Global analysis**

Based on the discussion in Chapters 2 and 3, the determinants and pathways of adverse birth outcomes are presumed to operate across both personal and contextual levels, and at different spatial scales, which require the use of specialized analytic tools for a comprehensive examination of association and causal relationships. Although it is almost universally agreed that multilevel modeling is the appropriate statistical modeling technique to test the social and environmental determinants of health inequalities, researchers have tended to use analytic tools that cannot adequately handle these hierarchical constructs. Among the 169 studies of income inequalities and population health reviewed by Wilkinson and Pickett (2006), only 43 used a multilevel design, while most of the other studies used ecological or personal level analyses, despite the possibilities of “ecological fallacy” and “atomistic fallacy” respectively.

An ecological study is one in which the unit of observation and analysis is conducted at the aggregate or group level rather than individual or personal level. It is therefore most appropriate in

explaining variation among groups or between areas when the constructs of interest can be conceptualized at an aggregated level. Sometimes personal-level data are not available and ecological studies are used to draw inferences on personal-level relations. This may cause an “ecological fallacy” because relationships observed between groups may not necessarily hold at the individual level. Similarly, an “atomistic fallacy” may occur if inferences regarding inter-group variability are incorrectly drawn from personal-level studies (Hox, 2002).

The existence of ecological and atomistic fallacies in and of themselves does not imply that a multilevel design is inherently superior as the fallacies are problems of inference, not of measurement. Hence, method selection depends on the nature and purposes of the questions under study. In this context, various theoretical, empirical and statistical reasons can be used to justify the use of a multilevel design for investigation of neighbourhood influences on causal pathways underlying adverse birth outcomes.

Theoretically, as described in the social ecological model in Chapter 2, social determinants of health inequalities operate at multiple levels, from personal risks through to community and society determinants. Group-level or neighbourhood-level risks or properties are associated with personal-level characteristics, but are not necessarily based on attributes of individual members (Luke, 2004). Some neighbourhood properties can be obtained by aggregating or averaging individual characteristics, such as average neighbourhood income, which represent the composition of individual members if no intra-neighbourhood correlation exists. Some are based on the collective relationships between members, such as the social networks within a neighbourhood. Others are characteristics of factors, such as national immigration or regional health care policies. Thus, the effects of neighbourhood relations on adverse birth outcomes may not be identifiable with individual risks. A multilevel design can effectively model the construction and interaction of aggregations and their members at different levels to distinguish their different effects on a dependent variable.

Empirically, the positivist tradition argues that through careful control over experimental and observational conditions, for example through the random selection of control and comparison groups, the impact of a certain determinant on health can be examined by simple single-level techniques. While this approach is effective in measuring the behaviours of closed systems in natural science research, it severely restricts the ability to measure or evaluate extra-individual, contextual effects of the social determinants of health. First, it is difficult to set up a case-control design in social science, since it is often impossible in reality to restrict the access to some resources for one group and allow access for others. Second, many health-related data are secondary and observational, derived from censuses, surveys, vital statistics, and health care administration databases. These data are not specifically designed for specific health studies and therefore neither incorporate randomization nor

do they control for confounding variables. Last, even though a specific design may be used in a health survey, since studies of health inequalities deal with complex open systems, it is almost impossible to control, restrict, or remove the effects of outside contextual influences (Fotheringham, et al., 2002).

Statistically, when dealing with sophisticated conceptual problems, social scientists tend to disaggregate higher-level information to the individual level and utilize traditional personal-level multiple regression methods in their analysis. This is problematic since the unmeasured contextual information is pooled into a single individual error term of the model. Contextual impacts mean that individuals within the same social context will have correlated errors, which violates a basic assumption of a multiple regression that the error term should be independent of the other variables. The estimated results may then deviate from the true relations. A model which ignores context also mistakenly implies that social processes behave in the same way in different contexts to influence individual characteristics. Thus, it is theoretically, empirically and statistically sound to use an appropriate multilevel design to conduct studies of adverse birth outcome inequalities.

#### 4.2.1 Multilevel modeling

The multilevel structuring or grouping of individuals may be non-spatial, for example, through different ethnicities, or spatial, by areas. A two-level spatial structure of level-1 units nested within level-2 neighbourhoods can be constructed, as shown in Figure 4.2. Because individual adverse birth outcomes are anticipated to depend on neighbourhoods, personal responses within a neighbourhood are likely to be similar.

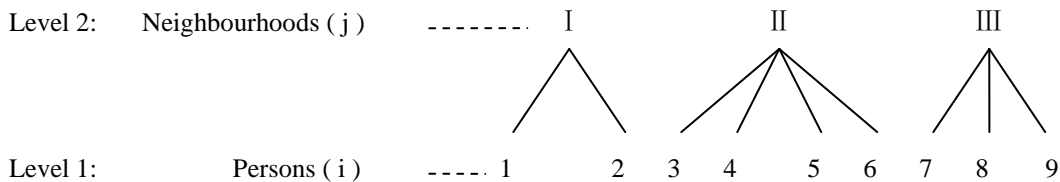


Figure 4.2 A two-level structure of persons nested in neighbourhoods

A traditional ordinary least squares (OLS) regression model (Wonnacott and Wonnacott, 1981) is constructed as:

$$Y_i = \beta_0 + \beta_1 X_i + e_i, \quad (4.1)$$

where  $Y$  is the response or dependent variable and  $X_i$  is the predictor or independent variable,  $\beta_0$  denotes the intercept coefficient and  $\beta_1$  denotes the slope coefficient of  $X_i$ . The  $e$  term is the independent random error term. This model conceptualizes that the predictor has a fixed linear effect on the response, and it operates only at one level. If  $Y$  denotes birth weight and  $X$  denotes smoking behaviour, this model suggests that smoking has a fixed effect on babies' birth weight everywhere or



for everyone regardless where people live or which group they belong to.

Because of the potential neighbourhood effects on both personal smoking behaviour and the chances of adverse birth outcomes, a multilevel design is more appropriate. A basic two-level linear regression model can then be written as (Luke, 2004):

$$\begin{aligned}\text{Level 1 (personal):} \quad & Y_{ij} = \beta_{0j} + \beta_{1j}X_{ij} + e_{ij} \\ \text{Level 2 (neighbourhood):} \quad & \beta_{0j} = \gamma_{00} + \gamma_{01}W_j + v_{0j} \\ & \beta_{1j} = \gamma_{10} + \gamma_{11}W_j + v_{1j}.\end{aligned}\tag{4.2}$$

These equations clearly delineate the multilevel nature of the model. The level 1 part is similar to the typical OLS model in Equation (4.1). However, instead of a single model, the  $j$  subscripts imply that for each level-2 unit or each neighbourhood,  $j$ , a different level-1 model is being estimated. In the above example, each neighbourhood may have a different average birth weight ( $\beta_{0j}$ ) and a different effect of smoking behaviour on birth weight ( $\beta_{1j}$ ). Thus, the multilevel model allows the intercept and slope to vary across the neighbourhoods. It conceptualizes that level-1 intercepts and slopes are the outcomes of level-2 predictors.

The two level-2 equations model the level-1 intercept and slope as functions of level-2 predictors and variability. In the first level-2 equation,  $\beta_{0j}$  is the level-1 intercept in level-2 unit  $j$ ;  $\gamma_{00}$  is the mean of the level-1 dependent variable, controlling for level-2 effects;  $\gamma_{01}$  is the effect of the level-2 predictor  $W_j$  on level-1 intercept;  $v_{0j}$  is the error or unmodeled variability for the level-1 intercept in unit  $j$ . Similarly, in the second equation,  $\beta_{1j}$  is the level-1 slope in level-2 unit  $j$ ; and  $\gamma_{10}$  is the mean of the level-1 slope; controlling for level-2 effects;  $\gamma_{11}$  is the effect of the level-2 predictor  $W_j$  on the level-1 slope; and  $v_{1j}$  is the unmodeled variability for the level-1 slope in unit  $j$ .

Substituting the level-2 equations into the level-1 equation, a single equation is obtained as:

$$Y_{ij} = [\gamma_{00} + \gamma_{10}X_{ij} + \gamma_{01}W_j + \gamma_{11}W_jX_{ij}] + [v_{0j} + v_{1j}X_{ij} + e_{ij}].\tag{4.3}$$

(Response) (Fixed/Average Parameters) (Random/Variance Parameters)

This model is compact, but the multilevel structure is hard to discern. On the other hand, the model clearly indicates which part is composed of fixed effects and which part is composed of random effects. It also shows that the level-1 parameters ( $\beta$ ) are not directly estimated, but are indirectly estimated through the level-2 parameters ( $\gamma$ ).

Equation (4.2) illustrates a typical two-level model with one predictor at level-1 and one predictor at level-2. However, there are many types of multilevel models that may be constructed and estimated depending on the research problem. For example, there may be more than 2 levels and there may be multiple predictors at each level. Intercepts or slopes may or may not predict the response, depending on one or both of them are at work in the context of the study. Hence, such multilevel

models can quickly become very complex to estimate as the number of levels and number of independent variables increase.

Three broad classes of multilevel models can be identified (Luke, 2004). The first type is the unconstrained model, which has no level-1 and level-2 predictors. This model describes both the intra-group and inter-group distributions of the response variables and calculates how much of the total variation can be attributed to individual variation and group variation respectively. In the above example, it would measure how birth weights are distributed between individuals within each neighbourhood and between neighbourhoods. Thus, the proportion of the birth weight variation accounted for by neighbourhoods or by individuals can be discerned. Hence, this model can be used as a starting point for building more complex models.

In contrast to a fixed intercept and fixed slope OLS model (Figure 4.3a), the second type of model assumes that level-1 intercepts vary across level-2 units, but that the level-1 slopes are fixed (Figure 4.3b). This type of model can be used if it is assumed, for example, that the smoking-disease relationship is the same in each neighbourhood, but some neighbourhoods have uniformly lower birth weights than do others, due to the impacts of neighbourhood characteristics.

The third type of model assumes that both intercepts and slopes vary across level-2 units. This model would be used under the assumption that there was a cross-level interaction between smoking behaviours and neighbourhood characteristics on birth weights. This interaction may take several forms. For example, in Figure 4.3c, the pattern is such that neighbourhood characteristics make very little difference for non-smokers, but there is a greater degree of neighbourhood variation in birth weights among the smokers. In contrast, Figure 4.3d shows relatively large neighbourhood differentials for non-smokers. Figure 4.3e shows that some neighbourhoods have greater impacts on birth weights for smokers and others have greater impacts for non-smokers. Finally, Figure 4.3f shows that there is no overall relationship between birth weights and smoking, but specific neighbourhoods may have distinctive relationships. Therefore, through different multilevel models, the spatial structures in which social determinants of health exert their impacts on adverse birth outcome inequalities can be considered. The pathways may be identified by examining how higher-level factors affect lower-level factors, which consequently affect personal birth outcomes.

In the analysis presented in this thesis, the dependent variables, namely IUGR and preterm births, are conventionally measured and have binary outcomes, where the value 1 represents having a LBW or preterm birth and the value 0 represents not having a LBW or preterm birth. These binary outcomes cannot be directly modeled by the above conventional multilevel linear models as they violate the normality and continuity assumptions of the response variables after controlling for covariates.

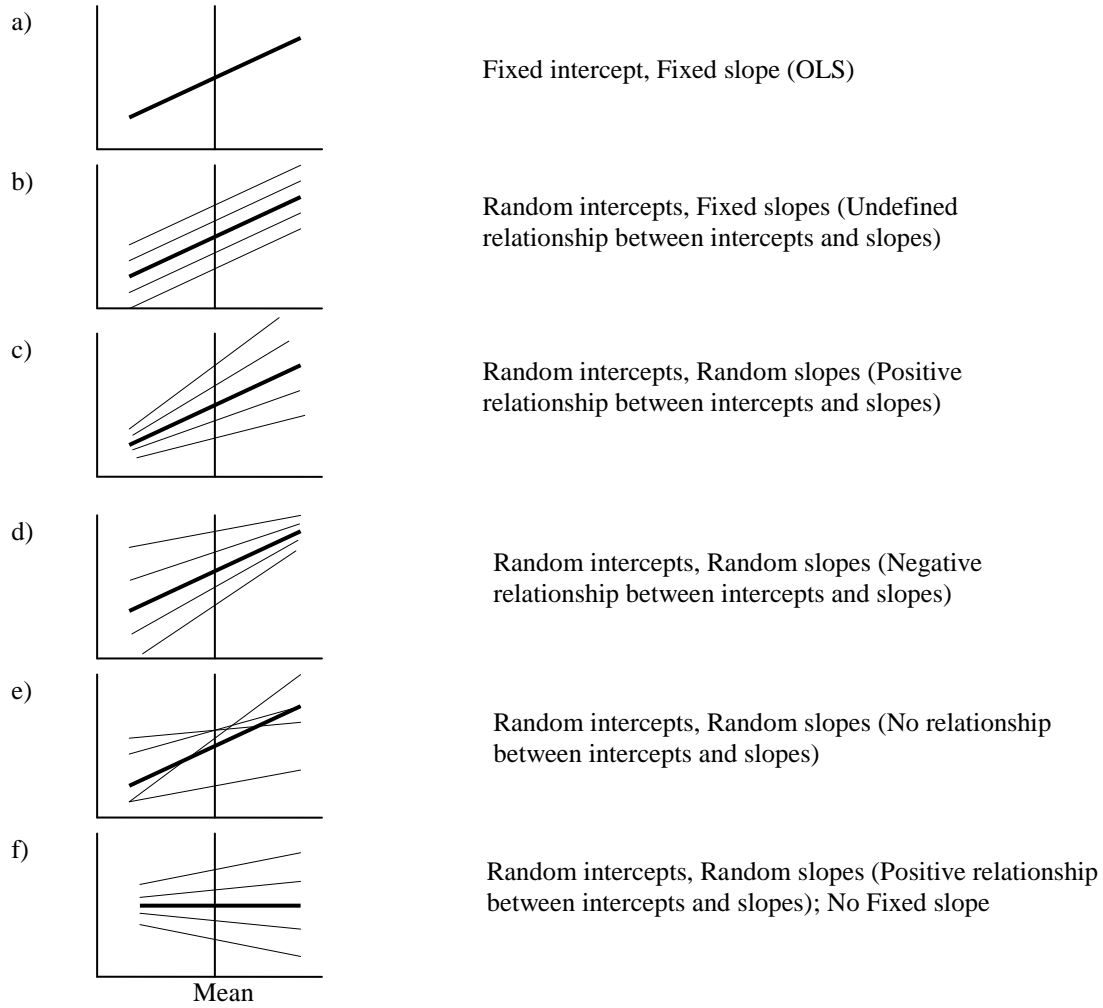


Figure 4.3: Varying level-1 and level-2 relationships (After Subramanian et al., 2003, p71)

However, with modification, the above models can be extended to handle this situation by using a generalized hierarchical linear model (GHLM), which works by including a necessary transformation and an appropriate error distribution for the dependent variable. Instead of using the binary outcomes directly, it is assumed that the underlying probability distributions for the binary outcomes are Bernoulli with estimated mean,  $p$ , which is interpreted as the probability of the event occurring. Hence, Equation (4.2) can then be modified as:

$$\begin{aligned}
 &\text{Adverse birth outcomes} \sim \text{binary}(p_{ij}) \\
 &\text{Level 1 (personal):} \quad \text{logit}(p_{ij}) = \beta_{0j} + \beta_{1j}X_{ij} \\
 &\text{Level 2 (neighbourhood):} \quad \beta_{0j} = \gamma_{00} + \gamma_{01}W_j + v_{0j} \\
 &\quad \quad \quad \beta_{1j} = \gamma_{10} + \gamma_{11}W_j + v_{1j}, \quad (4.4)
 \end{aligned}$$

where  $p_{ij}$  is the probability of having an adverse birth outcome for individual  $i$  in areal unit  $j$ , which follows a binary or Bernoulli distribution. There is no random term for the level-1 errors. The

variance for a binary variable is completely determined by the mean and thus is not a separate term to be estimated. Level-2 models are constructed as before. Instead of using  $p_{ij}$  directly, a link function  $\text{logit}(p_{ij}) = \log(p_{ij} / (1 - p_{ij}))$ , is used. The *Logit* function is a typical transformation for a binomial model, connecting the linear predictors to the mean of the outcome variable (LBW or preterm birth) and not directly to the outcome variable itself. All other terms remain the same.

The use of the *logit* function of adverse birth outcomes instead of the probability of adverse birth outcomes as the dependent variable is recommended since a linear specification of the probability of adverse birth outcome variable does not restrict the estimated values to lie within the limits [0, 1]. However, this restriction is crucial since the conditional expectation of adverse birth outcomes is interpreted as the probability of an event occurring given the value of all the explanatory variables. A linear specification becomes more problematic for data with clustering near the limits. The specification of the *logit* function overcomes this limitation by restricting the conditional expectation to lie within the range [0, 1]. Hence, using the GHLM models, different multilevel impacts of adverse birth outcomes can be modeled and tested accordingly.

As indicated in Chapter 2, IUGR and preterm births may reflect different intra-uterine and perinatal processes and experiences and therefore should be modeled separately. However, it is difficult to measure IUGR directly. Instead, IUGR is modeled by using LBW as the outcome while controlling for personal-level preterm births and multiple births. This specification allows singular full-term low birth weight births to be modeled and therefore represent IUGR, although this representation may be somewhat conservative. The age of mother and sex of baby are also controlled at the personal level so that the modeled covariates may not be confounded by these factors. Although mother's age is considered to be a modifiable risk factor, especially for teenage mothers, the objective is to establish the neighbourhood impacts on personal medical, social, psycho-social, and behavioural patterns. The risks associated with mother's age include both medical conditions (immaturity or associated chronic diseases) and behavioural impacts, which are confounded with the mother's other risk factors. By controlling for mother's age, the "pure" modifiable risk factors on IUGR and preterm births can then be identified.

A first pass examination of the associations of personal-level risks and IUGR or preterm births can then be constructed using the model:

*Adverse birth outcomes* ~ *binary* ( $p_{ij}$ )

$$\begin{aligned} \text{Level 1 (personal): } \text{logit}(p_{ij}) = & \beta_{0j} + \beta_1 \text{AGE19}_{ij} + \beta_2 \text{AGE36}_{ij} + \beta_3 \text{FEMALE}_{ij} + \beta_4 \text{MLTIBIRTH}_{ij} \\ & + (\beta_5 \text{PRETERMBIRTH}_{ij}) + \beta_6 \text{PERSONAL\_RISK}_{ij}, \end{aligned}$$

$$\text{Level 2 (neighbourhood): } \beta_{0j} = \gamma_{00} + v_{0j}. \quad (4.5)$$

A model to test neighbourhood-level risks can be similarly constructed as:

*Adverse birth outcomes* ~ binary ( $p_{ij}$ )

$$\text{Level 1 (personal): } \text{logit}(p_{ij}) = \beta_{0j} + \beta_1 \text{AGE19}_{ij} + \beta_2 \text{AGE36}_{ij} + \beta_3 \text{FEMALE}_{ij} + \beta_4 \text{MLTIBIRTH}_{ij} \\ + (\beta_5 \text{PRETERMBIRTH}_{ij}),$$

$$\text{Level 2 (neighbourhood): } \beta_{0j} = \gamma_{00} + \gamma_{01} \text{NB\_RISK}_j + v_{0j}, \quad (4.6)$$

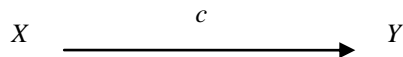
where *Adverse birth outcomes* can be either LBW or preterm births. *AGE19* is an indicator for a teenage mother ( $\leq 19$ ). *AGE36* is an indicator for a mother of advanced age ( $> 35$ ). *FEMALE* indicates a female baby. *MLTIBIRTH* is an indicator for multiple births. *PRETERMBIRTH* is an indicator for preterm birth, and  $(\beta_5 \text{PRETERMBIRTH})$  is only fitted in the model when LBW is the dependent variable. *PERSONAL\_RISK* represents a risk at personal level to be tested. *NB\_RISK* represents a neighbourhood-level risk. By fitting one risk factor at a time into these models, the associations between potential risk factors and IUGR or preterm births controlled for age of mother and sex of baby can be tested.

Similarly, once personal-level risks are identified or constructed to represent personal-level variations (a factor analysis is conducted on identified personal risks to construct distinct factors of personal risks, see Section 4.4.3 for details), Hypothesis 2 (compositional vs. contextual) can be tested by including these personal-level risk factors into the level-1 model in Equation (4.6). After controlling for personal risk factors, if significant associations still remain between any neighbourhood-level risks and adverse birth outcomes, contextual impacts can then be confirmed. The hypothesized pathways models (Hypotheses 3 and 4) can then be tested on the associated neighbourhood risks using a mediational analysis. This is explained in the following section.

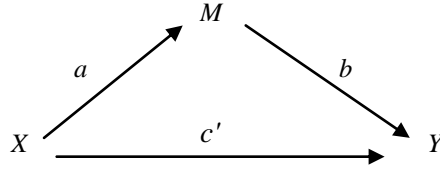
#### 4.2.2 Mediational analysis

A mediational model (MacKinnon et al., 2007) is often used to test potential causal pathways between exposures and responses. Based on presumed causal pathways, a mediational model uses process analysis to test the mechanism by which the initial variable affects the outcome through the mediator.

If an initial variable  $X$  is assumed to affect the outcome  $Y$ , the unmediated direct relation can be modeled as:



If the effect of  $X$  on  $Y$  is also assumed to be mediated by a mediating or intervening variable  $M$ , the mediated model is:



Path  $c$  in the unmediated model is called the total effect. In the mediated model, the variable  $X$  may still affect  $Y$  through path  $c'$  which is called the direct effect. Complete mediation occurs when variable  $X$  no longer directly affects  $Y$  after  $M$  has been controlled and path  $c'$  is zero. Partial mediation occurs when the path from  $X$  to  $Y$  is reduced but is still different from zero when the mediator is controlled.

Baron and Kenny (1986) and Judd and Kenny (1981) have discussed four steps in establishing mediation:

Step 1: Test that  $X$  is correlated with  $Y$  (path  $c$ ) to establish that there is an effect that may be mediated.

Step 2: Test that  $X$  is correlated with  $M$  (path  $a$ ). This step involves treating  $M$  as if it were an outcome of  $X$ .

Step 3: Test that  $M$  affects  $Y$  (path  $b$ ). It is not sufficient just to correlate  $M$  with  $Y$ , since  $M$  and  $Y$  may be correlated because they are both caused by the initial variable  $X$ . Therefore,  $X$  must be controlled in establishing the effect of the mediator on the outcome. A multivariable regression or multilevel regression can be used to estimate  $Y$  as dependent on both  $X$  and  $M$ .

Step 4: Use the same regression equation as step 3. If the significance testing shows that the effect of  $X$  on  $Y$  controlling for  $M$  (path  $c'$ ) is zero,  $M$  completely mediates the  $X$ - $Y$  relationship. Otherwise  $M$  only partly mediates the relationship.

When the variables  $M$  and  $Y$  are continuous, so that the regression models are linear, the amount of mediation, which is called the indirect effect, is measured by the product of the effect of  $X$  on  $M$  times the effect of  $M$  on  $Y$ , or  $ab$ . The total effect  $c$  is the sum of direct effect  $c'$  and indirect effect  $ab$ . A Sobel test is usually used for testing the significance of  $ab$  (Sobel, 1982). In this test, the standard error of  $ab$  is estimated as:

$$b^2 s_a^2 + a^2 s_b^2, \quad (4.7)$$

where  $s_a$  and  $s_b$  are the standard errors of  $a$  and  $b$  respectively. A  $z$ -test can then be conducted in order to compute the statistical significance of the mediating effect.

As expressed in Steps 1-4, the procedures to test a mediational effect are not model specific. In other words, they are not dependent on the actual statistical model, whether it be a least squares regression model, a multilevel model, or a Bayesian model. The statistical models created above can

therefore all be decomposed to test the hypothesized pathways and the mediational effects of personal-level risks on the impact of contextual-level risk on adverse birth outcomes. In this thesis, the outcome (IUGR or preterm births) is binary and the statistical model is multilevel. This makes the analysis more complex than for OLS models. Table 4.1 shows the comparison of OLS and Multilevel expressions for the three mediational analysis equations.

<i>Relation</i>	<i>OLS</i>	<i>Multilevel</i>
XY relation (path c)	$Y = k + cX + e$	$Y_{ij} \sim \text{binary}(p_{ij})$ Personal level: $\text{logit}(p_{ij}) = \alpha_{0j}$ Area level: $\alpha_{0j} = \gamma_{00} + cX_j + v_{0j}$
MX relation (path a)	$M = k' + aX + e'$	Personal level: $M_{ij} = \alpha'_{0j} + e_{ij}$ Area level: $\alpha'_{0j} = \gamma'_{00} + aX_j + v'_{0j}$
XMY relation ( path b, c')	$Y = k'' + c'X + bM + e''$	$Y_{ij} \sim \text{binary}(p_{ij})$ Personal level: $\text{logit}(p_{ij}) = \alpha''_{0j} + bM_{ij}$ Area level: $\alpha''_{0j} = \gamma''_{00} + c'X_j + v''_{0j}$

Table 4.1: OLS and Multilevel binary outcome model comparison for mediational analysis

Although Table 4.1 shows that each OLS equation for mediational analysis has its corresponding equations for the multilevel binary outcome model, some of the relationships found in OLS mediational models do not hold for multilevel models. The two estimates of the mediated effect,  $ab$  and  $c - c'$  are algebraically equivalent in OLS models, but the equivalence does not hold in the multilevel framework due to the nature of multilevel estimation. Estimates of direct and indirect effects will not necessarily sum to the estimates of the total effect if estimates of the different effects are generated using different multilevel equations. In addition, a logistic multilevel regression includes no personal-level random effects. This leads to the estimated probability values having different variance scales for the X-Y relation and the X-M-Y relation, which make the comparison between equations more difficult.

A simulation study conducted by Krull and MacKinnon (1999) suggested that for normal output multilevel models, although  $c - c'$  and  $ab$  estimates of the mediated effects are not algebraically equivalent in multilevel models, the difference can be ignored based on the simulation study results. For binary logistic model comparisons, MacKinnon and Dwyer (1993) also proposed a method to make the variance equivalent across equations by standardizing regression coefficients prior to estimating mediation. However, experiments show that this method is not so effective in multilevel cases.

A series of analysis procedures is therefore proposed to make a mediational analysis on

multilevel binary outcome models possible and put the equations on the same scales so that comparison of  $c - c'$  and  $ab$  is possible. First, suppose the mediator  $M$  is a normally distributed personal level variable.  $X$  is a group level explanatory variable that is supposed to be mediated by  $M$ . Suppose  $Y$  is the response variable and it is binary (1 – has condition of interest, 0 – otherwise). To test the association between  $X$  and  $M$ ,  $M$  is regressed on  $X$  using the following equation,

$$M = aX + U, \quad (4.8)$$

where  $a$  is the fixed effect for  $X$ , and  $U$  is a random intercept (intercept + random effect). This model partitions  $M$  into a  $X$ -correlated part and a  $X$ -non-correlated part. It assumes that the individual error in the prediction of  $M$  by  $X$  is negligible in comparison with the group level error  $U$ .

The  $X$ - $M$ - $Y$  associations can be modeled as follows.

$$Y \sim \text{Binary}(p)$$

$$Y^* = \text{Logit}(p) = c'X + bM + \gamma Z,$$

where  $Z$  is the group-level random effect.

Substituting  $M$  from Equation (4.8), the combined model is:

$$Y^* = \text{Logit}(p) = c'X + b(aX + U) + \gamma Z$$

$$= (c' + ab)X + bU + \gamma Z. \quad (4.9)$$

Since  $X$ ,  $U$ , and  $Z$  are uncorrelated, the variance of the predicted log odds is:

$$\text{Variance}(Y^*) = \text{variance}((c' + ab)X) + \text{variance}(bU) + \text{variance-covariance}(\gamma Z).$$

Thus, the variance of the predicted value is explained by  $X$ ,  $U$  (the part of  $M$  that is uncorrelated with  $X$ ), and the group level variance. For the part that is predicted by  $X$ ,  $c'$  is the direct effect, and  $ab$  is the indirect effect.  $Y^*$  can then be seen as a continuous variable and can be regressed to  $X$  to obtain the total effect:

$$Y^* = cX + \gamma_I Z_I + e. \quad (4.10)$$

In this formulation, the variance of  $Y^*$  is explained by the variance of  $cX$ , the variance of  $e$ , and the group level variance of  $\gamma_I Z_I$ . Comparing it with the above combined model, the variance of  $bU$  is approximated by the variance of  $e$ , and the variance-covariance of  $\gamma Z$  is approximated by the variance of  $\gamma_I Z_I$ . The coefficient  $c$  can then be seen as the total effect of  $X$  on  $Y$ . In the case of a single-dimensional  $M$ ,  $c$  should be approximately equal to  $c' + ab$ . In the case of a multidimensional  $M$  with orthogonal components,  $c$  should be approximately equal to  $c' + \text{the sum of the } ab \text{ terms}$  corresponding to the dimensions of  $M$ . Assuming the mediators are orthogonal, the standard error of the overall indirect effect (*the sum of the  $ab$  terms*) can also be approximated by



$$\sqrt{\sum_{i=1}^n sd(a_i b_i)^2},$$

where  $sd(a_i b_i)$  is the standard error of  $a_i b_i$  estimated using Equation (4.7), and  $n$  is the number of mediators. However, this equation should be used with caution especially when sample sizes are small and the underlying normality assumption for the parameter distributions may be violated.

Using this mediational analysis procedure, Hypotheses 3 and 4 can then be tested and the pathways of each neighbourhood risk through personal risks to adverse birth outcomes can be identified. Furthermore, the mechanisms of neighbourhood influences on the dependent variable can be revealed.

#### 4.2.3 Testing of neighbourhood variations

Before pathway models can be tested, it is important to examine spatial or neighbourhood impacts of adverse birth outcomes. Spatial impacts are multi-dimensional, including both global and local aspects. The global aspect of spatial impacts involves testing the first hypothesis, which, as discussed in Chapter 3, includes both non-spatial and spatial considerations. First, non-spatial tests assume spatial independence between neighbourhoods. Based on this assumption, significant global variations may exist from neighbourhood to neighbourhood, compared with individual variations. In other words, the occurrence of individual incidences may be correlated within each neighbourhood (due to potential neighbourhood risks), but uncorrelated between neighbourhoods, so that high occurrence rates can be observed in some neighbourhoods while low occurrence rates exist in others. Second, spatial tests assume that general spatial clustering or autocorrelation of adverse birth outcomes may be apparent in the study region. In contrast with the non-spatial tests, they assume that spatial autocorrelation exists among nearby neighbourhoods. The assessment of global clustering is intended to test the existence of this overall aspect of the clustering tendency.

Local assessment of spatial impacts involves tests of “hot spot” clustering of adverse birth outcomes and spatial variations of the impacts of risks. Methods of local spatial testing are discussed in Section 4.3, whereas global tests on non-spatial neighbourhood-level variation and spatial clustering are now considered.

##### 4.2.3.i Intraclass correlation coefficient

The intraclass correlation coefficient (ICC) provides a quantitative measure of similarity between individuals within groups or within areal units. It can be used to quantify the extent of neighbourhood variations of adverse birth outcomes. Suppose that there are  $k$  areal units. The  $i^{th}$  areal unit contains  $n_i$  individuals, each having a binary response  $X_{ij}$  (1 - Success, 0-Failure). Let  $Y_i = \sum_j X_{ij}$  represent the total number of successes in the  $i^{th}$  areal unit. The underlying model for the probability

of success is assumed to be the same for all individuals irrespective of the individual's areal unit, namely  $P(X_{ij}=1) = \pi$  for all  $i, j$ . It also assumes that the responses from different areal units are independent while the within group correlation of any two responses is  $\rho$ . The correlation is assumed not to vary with group size. Following this assumption, the areal level variation can be defined as:

$$\text{Var}(Y_i) = n_i \pi (1 - \pi) [1 + (n_i - 1) \rho].$$

The areal-level variation is the sum of individual variations if  $\rho = 0$  and it is overdispersed relative to the binomial distribution if  $\rho > 0$  and underdispersed if  $\rho < 0$ . Under this assumption, various statistical methods have been developed to calculate the ICC.

One method developed by Fleiss and Cuzick (1979) is through an estimator with a direct probabilistic interpretation, namely

$$\rho_{FC} = 1 - \frac{1}{(N - K)\hat{\pi}(1 - \hat{\pi})} \sum_{i=1}^k \frac{Y_i(n_i - Y_i)}{n_i}, \quad (4.11)$$

where  $\rho_{FC}$  is the estimated intraclass correlation coefficient (ICC),  $N$  is the total number of observations,  $k$  is the number of groups or areal units,  $\hat{\pi} = \sum Y_i / \sum n_i$  is the overall proportion of successes,  $Y_i = \sum_j X_{ij}$  denotes the total number of successes in the  $i^{th}$  group,  $X_{ij}$  is a binary response indicating the success of the  $j^{th}$  individual in the  $i^{th}$  group, and  $n_i$  is the number of individuals in the  $i^{th}$  group.

A second method used to calculate the ICC is through an estimator based on direct calculation of the correlation within each group. The intuitive method of estimating intraclass correlation is simply to calculate the Pearson correlation coefficient over all possible pairs of observations within groups. However, this method tends to give a much higher weight to large groups. Karlin, et al. (1981) have proposed a more general weighted estimator for binary data to account for this problem, namely

$$\rho = \frac{\sum_{i=1}^k w_i Y_i (Y_i - 1) - \hat{\mu}^2}{\hat{\mu}(1 - \hat{\mu})},$$

where  $\hat{\mu} = \sum_{i=1}^k w_i (n_i - 1) \sum_{j=1}^{n_i} X_{ij}$  and the weights  $w_i$  satisfy  $\sum_{i=1}^k n_i (n_i - 1) w_i = 1$ . Based on different forms of  $w_i$ , the estimator can be different. Taking  $w_i = 1/[N(n_i - 1)]$ , the estimated mean is  $\hat{\mu} = \frac{1}{N} \sum_{i=1}^k Y_i$  and the Pearson intra-class correlation becomes,

$$\rho_{PPR} = \frac{1}{\hat{\mu}(1 - \hat{\mu})} \left[ \frac{1}{N} \sum_{i=1}^k \frac{Y_i(Y_i - 1)}{n_i - 1} - \hat{\mu}^2 \right]. \quad (4.12)$$

A third method is through the analysis of variance (ANOVA) estimator (Fleiss, 1981), which is

given by

$$\rho_{AOV} = \frac{BMS - WMS}{BMS + (n_A - 1)WMS} , \quad (4.13)$$

where  $n_A = \frac{1}{k-1} [N - \sum_{i=1}^k \frac{n_i^2}{N}]$  .  $BMS = \frac{1}{k-1} [\sum_{i=1}^k \frac{Y_i^2}{n_i} - \frac{1}{N} (\sum_{i=1}^k Y_i)^2]$  and  $WMS = \frac{1}{N-k} [\sum_{i=1}^k Y_i - \sum_{i=1}^k \frac{Y_i^2}{n_i}]$  are between-cluster and within-cluster mean squares respectively.

A further confidence interval estimation of this estimator was also carried out (Lui et al., 1996). Let us define random variables  $B_i = n_i (\frac{Y_i}{n_i} - \pi)^2$ , where  $\pi$  is the probability of success for the conditional binomial distribution of  $Y_i$ , and  $W_i = \sum_j (X_{ij} - Y_i)^2 / (\mu_c - 1)$ , where  $\mu_c$  is the expected value of  $n_i$ . The expectation  $E(B_i)/E(W_i)$  is assumed to be equal to a constant  $\theta$ . Thus, the random variable  $Z_i = B_i - \theta W_i$  has 0 mean and constant variance  $var(Z_i)$ . By the Central Limit Theorem, as the number of groups or areal units  $k$  increases, the random variable  $\sqrt{k} \sum_{i=1}^k Z_i / k$  asymptotically has a normal distribution with mean 0 and constant variance  $var(Z_i)$ . Using the average of  $Y_i$  and the average of  $n_i$  to replace  $\pi$  and  $\mu_c$  respectively, then  $\frac{\sqrt{k} \sum_{i=1}^k Z_i}{k}$  can be rewritten as  $[(k-1)/k] BMS - \theta WMS$ . For large  $k$ , the probability  $P((BMS - \theta WMS)^2 / (var(Z)/k)) \leq Z_{\alpha/2}^2)$  is approximately equal to  $1-\alpha$ , where  $BMS^* = ((k-1)/k) BMS$  and  $Z_\alpha$  is the upper  $100\alpha$ th percentile of the standard normal distribution. Solving the implied quadratic equation gives an approximate  $1-\alpha$  confidence interval for the intraclass correlation with lower limit  $(\theta_l - 1) / (\theta_l + n - 1)$  and upper limit  $(\theta_u - 1) / (\theta_u + n - 1)$ , where  $\theta_l$  and  $\theta_u$  are the two distinct real roots of the quadratic equation in  $\theta$ :  $(BMS^* - \theta WMS)^2 / var(Z)/k = Z_{\alpha/2}^2$ .

Other methods, such as the moment estimators (Yamamoto and Yanagimoto, 1997) and extended quasi-likelihood and pseudo-likelihood estimators (Nelder and Pregibon, 1987), have also been developed to calculate the ICC. However, all the above described methods tend to underestimate the intraclass variation when  $\rho$  is small, and  $\rho$  can be negative if the observed distribution is under-dispersed relative to the assumed binomial distribution (Ridout et al., 1999). In this case, it is difficult to interpret ICC as the proportion of total variance.

More recently, under the framework of random effects models, an alternative approach that has gained attention is to use a generalized linear mixed model in which the effects of the covariates are assumed to be linear on a transformed (e.g., logit or probit) scale, and a random group effect (or effects) is added on the transformed scale. For example, adapted from Equation (4.4), a probability model for clustered binary outcomes with an explanatory indicator  $X_{ij}$  can be constructed as:

$$\begin{aligned}
& \text{Adverse birth outcomes} \sim \text{binary} (p_{ij}) \\
& \text{Level 1 (personal):} \quad \text{logit}(p_{ij}) = \beta_{0j} + \beta_1 X_{ij} \\
& \text{Level 2 (neighbourhood):} \quad \beta_{0j} = \gamma_{00} + v_{0j}.
\end{aligned} \tag{4.14}$$

In this equation,  $u_{0j}$  is assumed to be stochastically independent of the covariate  $X_{ij}$  and normally distributed with mean 0 and unknown variance  $\sigma^2$ . This model is regarded as a random effects logistic probability model or a mixed effects logistic probability model indicating that the log-odds of success ( $Y_{ij}=1$ ) is a function of both a fixed effect and a random effect (Zeger et al., 1988; Ten Have et al., 2004). Under this model, the variance of the random effect  $v_{0j}$ ,  $\sigma^2$ , represents the between-areal unit variation. The within-areal unit variation is  $\pi^2/3$ , representing the variance of the fixed effects logistic distribution. Hence, the ICC can be written as:

$$ICC = \sigma^2 / (\sigma^2 + \pi^2/3). \tag{4.15}$$

Alternatively, a *probit* model can be fitted to estimate the ICC. The model is constructed similarly to the above *logit* model in Equation (4.14), except for the introduction of a standard normal random effect at the personal level. The *probit* model can be specified as:

$$\begin{aligned}
& \text{Adverse birth outcomes} \sim \text{binary} (p_{ij}) \\
& \text{Level 1 (personal):} \quad \text{Probit} (p_{ij}) = \beta_{0j} + \beta_1 X_{ij} \\
& \text{Level 2 (neighbourhood):} \quad \beta_{0j} = \gamma_{00} + v_{0j},
\end{aligned} \tag{4.16}$$

where the link function  $\text{Probit}(p_{ij}) = \Phi^{-1}(p_{ij})$  is the inverse standard normal cumulative distribution function and  $u_{0j}$  is normally distributed with mean 0 and unknown variance  $\sigma^2$ . Although not explicitly specified in the model, a  $N(0, 1)$  term is included at the personal level. Under this model, the areal-level variance of the random effect  $v_{0j}$  is  $\sigma^2$  and the personal level variance is 1. Thus, the ICC can be written as:

$$ICC = \sigma^2 / (\sigma^2 + 1). \tag{4.17}$$

The ICC estimators constructed by the random effect models are always non-negative, allowing them to be interpreted as the proportion of total variance that is between neighbourhoods. The other benefit of these ICC estimators is that they can be generalized to allow for hierarchical multi-covariate effects, in which case the ICC is interpreted as capturing the within-areal unit similarity of the covariate-adjusted outcomes. However, a limitation of all above ICC estimators is that they only measure the proportion of neighbourhood variation to the total variations, and do not account for the presence of spatial autocorrelation between the neighbourhood areal units. In certain circumstances, even though the neighbourhood variation is small, considerable spatial clustering may exist due to the impacts of neighbourhood risks. A further global spatial clustering analysis can be carried out to

address this issue and find out if there is a recognizable spatial clustering of adverse birth outcomes.

#### 4.2.3.ii Global spatial clustering

Spatial dependencies have many different dimensions and can take various forms. Tests of spatial clustering can be classified as focused tests, which assess the clustering of incidence around a pre-fixed point of potential source of risk, and general tests, which investigate whether clustering occurs over an entire study region. The second group of tests can be further classified as global spatial clustering tests for examining a tendency to cluster, and local spatial clustering tests exclusively for searching for the location of clusters. This thesis focuses on general tests to examine whether or not global spatial clustering of adverse outcomes exists, and where hot spots of local significantly higher incidences of adverse birth outcomes are observed. The tests for general global clustering are discussed in this section and some of the local spatial clustering tests are reviewed and discussed in the following section.

Various general global clustering indices have been proposed and used in the literature (for example, Moran, 1948; Geary, 1954; Oden, 1995; Tango, 1995; Turnbull et al., 1990; Kulldorff and Nagarwalla, 1995; Besag and Newell, 1991). Many global indices of spatial autocorrelation have a similar form, often referred to as general cross-product statistics (Mantel, 1967). These indices are constructed as follows:

$$\frac{\sum_{i=1}^N \sum_{j=1}^N w_{ij} Sim_{ij}}{\sum_{i=1}^N \sum_{j=1}^N w_{ij}},$$

where  $N$  is the total number of areal units in the study region,  $Sim_{ij}$  is the similarity between data values at locations  $i$  and  $j$ , and  $w_{ij}$  denotes a weight describing the proximity between locations  $i$  and  $j$ , for  $i$  and  $j = 1, \dots, N$  and  $i \neq j$ . The collection of  $w_{ij}$  constructs a spatial proximity matrix or spatial connectivity matrix. The weight  $w_{ij}$  is the  $(i, j)^{th}$  element in the matrix and quantifies the spatial dependence between areal units  $i$  and  $j$ .

The weight  $w_{ij}$  can have various forms depending on the spatial layout of neighbourhoods. The simplest one is the spatial adjacency or contiguity matrix defined as:

$$w_{ij} = \begin{cases} 1 & \text{if } i \text{ and } j \text{ share a common boundary} \\ 0 & \text{otherwise.} \end{cases}$$

This can be expanded as:

$$w_{ij} = \begin{cases} 1 & \text{if } j \text{ is one of the } q \text{ nearest to } i \\ 0 & \text{otherwise.} \end{cases}$$

Instead of specifying the weights in terms of neighbourhoods,  $w_{ij}$  can also be defined as a

function of distance, such as:

$$w_{ij} = \begin{cases} 1 & \text{if } d_{ij} < \delta \\ 0 & \text{otherwise} \end{cases}$$

or

$$w_{ij} = \begin{cases} d_{ij}^\alpha & \text{if } d_{ij} < \delta \\ 0 & \text{otherwise,} \end{cases}$$

where  $d_{ij}$  is the distance between the centroids of areal unit  $i$  and  $j$ , and  $\alpha < 0$ .

This can also be defined as the fraction of  $i$ 's border shared with  $j$

$$w_{ij} = \begin{cases} \frac{l_{ij}}{l_i} & \text{if } i \text{ and } j \text{ share a common boundary} \\ 0 & \text{otherwise,} \end{cases}$$

where  $l_{ij}$  is the length of the common boundary between unit  $i$  and  $j$  and  $l_i$  is the perimeter of region  $i$ .

Different measures of similarity,  $Sim_{ij}$ , define different index classes. Two of the most commonly used global indices of spatial autocorrelation are Moran's  $I$  (Moran, 1950) and Geary's  $C$  (Geary, 1954).

Moran's  $I$  defines the similarity as the product of the respective difference between  $X_i$  and  $X_j$  with the overall mean:

$$sim_{ij} = (x_i - \hat{x})(x_j - \hat{x}),$$

where  $\hat{x}$  is the mean value of sample  $x_i$  in the study region. Dividing the basic form by the sample variance, yields:

$$I = \frac{n}{\sum_{i=1}^n \sum_{j=1}^n w_{ij}} \frac{\sum_{i=1}^n \sum_{j=1}^n w_{ij} (x_i - \hat{x})(x_j - \hat{x})}{\sum_{i=1}^n (x_i - \hat{x})^2}. \quad (4.18)$$

Moran's  $I$  evaluates whether the pattern expressed is clustered, dispersed, or random. In general, a Moran's  $I$  index value near 1.0 indicates perfect clustering while an index value near -1.0 indicates perfect dispersion. A value close to 0 indicates a randomized distribution.

Geary's  $C$  defines the similarity as:

$$sim_{ij} = (x_i - x_j)^2.$$

By scaling this basic form by the overall variation around the mean regional observation  $X$ , this can be defined as:

$$C = \left( \frac{N}{2 \sum_{i=1}^N (x_i - \hat{x})^2} \right) \frac{\sum_{i=1}^N \sum_{j=1}^N w_{ij} (x_i - x_j)^2}{\sum_{i=1}^N \sum_{j=1}^N w_{ij}}. \quad (4.19)$$

In contrast to Moran's  $I$ , low values of Geary's  $C$  denote spatial clustering and high values indicate a dispersed spatial distribution. It ranges in value from 0 to 2, with 0 indicating perfect positive spatial correlation and 2 indicating perfect negative spatial autocorrelation. The expected value of this index is 1 under the null hypothesis of spatial independence.

The purpose of the global clustering test used in this thesis is to prove the existence of significant overall spatial dependency so that further spatial analysis on the correlations of neighbourhood exposures and adverse birth outcomes can be warranted. No substantial examination of the extent to which global clustering may exist is needed. Thus, the classic global Moran's  $I$  test is used to examine the spatial dependency of adverse birth outcomes, instead of further in-depth exploration of different dimensions of spatial dependency. To test the statistical significance of Moran's  $I$  value, the  $z_I$ -score is computed as:

$$z_I = \frac{I - E[I]}{\sqrt{V[I]}},$$

where  $E[I] = -1/(n-1)$  and  $V[I] = E[I^2] - E[I]^2$ . The  $z$  score and associated  $p$ -value based on a standard normal distribution can then be used to test for statistical significance. When the  $p$ -value is small and the absolute value of the  $z$  score is large enough that it falls outside of the desired confidence interval, the null hypothesis that "there is no spatial clustering of the observed values in the study area" can be rejected. In other words, any observed spatial autocorrelation is true and unlikely caused by chance.

The indices of general cross-product statistics are criticized for not properly adjusting for a heterogeneous population density (Besag and Newell, 1991; Marshall, 1991). To account for this, a two-step procedure is used in this thesis to model the spatial dependency of adverse birth outcomes. Instead of directly measuring the Moran's  $I$  score of the raw adverse birth outcome rates, a model similar to Equation (4.5) is fitted and the residual or neighbourhood level random effect  $v_{0j}$  is obtained after controlling for personal-level risks. If there is no spatial dependency,  $v_{0j}$  should be randomly distributed over space. Otherwise, spatial autocorrelation may exist due to some unknown neighbourhood-level impacts. A Moran's  $I$  test can then be used to test this spatial dependency. Using the ICC estimators presented above and the global clustering test (Moran's  $I$ ), Hypothesis 1 stated in Chapter 3 can be effectively tested.

### 4.3 Local spatial analysis

While the ICC estimators and the global clustering tests provide methods to estimate the existence of the overall degree of spatial impacts, they do not provide a necessary means to identify the locations of spatial clustering or high incidences of adverse birth outcomes and test their significance. The possibility of spatial heterogeneity suggests that the estimated degree of

autocorrelation may vary significantly across geographic space. Locating local variations or “hot spots” of adverse birth outcomes is unlikely to provide clues about any specific aetiology that may be responsible for observed effects. However, it can provide clues about the underlying explanatory factors. Neighbourhoods detected with higher risks will also be good places to conduct further case-control or cohort studies. Practically, the identified high risk areas will be important places for public health surveillance and health-related resource allocations. They may also be the places to initiate community-oriented health programs and local multi-stakeholder participation to address local health needs.

In addition, the complexity of the personal, social and spatial determinants of population health in general and adverse birth outcomes in particular makes the causal relationships of identified risks and responses non-stationary over space. In contrast to the natural sciences, where a global statement of relationships that is suitable for every situation may be sought, social processes tend to be context-oriented and the impacts of risks may differ from place to place. Thus, a fixed parameterization of exposure-response associations may not be sufficient to explain the non-stationary nature of these relationships. Methods that can specify explicitly spatial variations of exposure-response associations should be sought. Methods are now discussed to assess the spatial variation of disease and the spatial variation of associations respectively in order test the last hypothesis stated in Chapter 3.

#### 4.3.1 Local clustering analysis

Local spatial clustering statistics provide estimates disaggregated to the level of the spatial analysis units, allowing assessment of the dependency relationships across space. Various local clustering analysis methods have been developed to obtain the location of clusters.

##### 4.3.1.i Local indicators of spatial association

The above mentioned global spatial autocorrelation statistics, such as Moran’s  $I$  and Geary’s  $C$ , all have their local form, which is often referred to as local indicators of spatial association (LISAs) (Anselin, 1995). Most LISAs are defined as local versions of the well-known global indices. The basic form of the LISA for areal unit  $i$  is defined as

$$\sum_{j=1}^N w_{ij} sim_{ij}.$$

The local Moran’s  $I$  is then given as:

$$I_i = \frac{x_i - \hat{x}}{S_i^2} \sum_{j=1, j \neq i}^n w_{ij} (x_j - \hat{x}), \quad (4.20)$$



where  $S_i^2 = \frac{\sum_{j=1, j \neq i}^n w_{ij}}{n-1} - \hat{x}^2$ , and all other notation is the same as the global Moran's  $I$ . The  $z$  scores of  $I_i$  are computed as:

$$z_{I_i} = \frac{I_i - E[I_i]}{\sqrt{V[I_i]}},$$

where  $E[I_i] = -\sum_{j=1, j \neq i}^n w_{ij} / (n-1)$  and  $V[I_i] = E[I_i^2] - E[I_i]^2$ . Statistical tests for significant local clustering can then be conducted accordingly.

The local version of Geary's  $C$  is calculated as:

$$C_i = \frac{n \sum_{j=1}^n w_{ij} (x_i - x_j)^2}{\sum_{j=1}^n (x_j - \hat{x})^2}. \quad (4.21)$$

The  $z$  scores of  $C_i$  are computed as

$$z_{I_i} = \frac{C_i - E[C_i]}{\sqrt{V[C_i]}},$$

where  $E[C_i] = \frac{2n}{n-1} \sum_{j=1, j \neq i}^n w_{ij}$  and  $V[C_i] = E[C_i^2] - E[C_i]^2$ .

The local Getis  $G_i^*$  statistic (Getis and Ord, 1992) is also developed in the form

$$G_i^* = \frac{\sum_{j=1}^n w_{ij} x_j - \hat{x} \sum_{j=1}^n w_{ij}}{S \sqrt{\frac{\sum_{j=1}^n w_{ij}^2 - (\sum_{j=1}^n w_{ij})^2}{n-1}}}, \quad (4.22)$$

where  $S = \sqrt{\frac{\sum_{j=1}^n x_j^2}{n} - (\hat{x})^2}$ . The  $G_i^*$  score itself is the  $z$  score and can be used for statistical testing under the standard normal distribution.

While the LISA statistics capture local variations, they do not take into account the hierarchical nature of space. The use of the direct counts or raw rates of diseases for calculation also fail to control for other confounding factors. To overcome these limitations, multilevel models with spatially structured random effects have been proposed.

#### 4.3.1.ii Spatially structured random effect models

The random effect model in Equation (4.14) can be modified to map the spatial variations of diseases to overcome these limitations. Instead of using the independent random effect  $v_{0j}$  in Equation (4.14), a spatially dependent random effect  $u_{0j}$  can be used to model the spatial autocorrelation. The

equation is then specified as

$$\begin{aligned}
& \text{Adverse birth outcomes} \sim \text{binary}(p_{ij}) \\
& \text{Level 1 (personal):} \quad \text{logit}(p_{ij}) = \beta_{0j} + \beta_1 X_{ij} \\
& \text{Level 2 (neighbourhood):} \quad \beta_{0j} = \gamma_{00} + u_{0j},
\end{aligned} \tag{4.23}$$

where  $u_{0j}$  is normally distributed with 0 mean and a variance-covariance matrix that incorporates spatial relationships among the spatial units used. Specifically, using standard multivariate normal theory (Besag and Kooperberg, 1995), the joint multivariate Gaussian model can be expressed as

$$\beta_{0j} \sim \text{Normal}(\gamma_{00}, v\mathbf{\Sigma} = v(I - \phi C)^{-1}M).$$

This model specification is called a Conditional Autoregressive (CAR) model (Cressie, 1993). In the model,  $C$  is a matrix with elements  $c_{ij}$  measuring the spatial association between  $i$  and  $j$ ,  $\phi$  is a parameter measuring spatial dependence, and  $M$  is a diagonal matrix with diagonal elements  $m_{jj}$  chosen to make the variance-covariance matrix  $\mathbf{\Sigma}$  symmetric and positive-definite. The parameter  $v$  is a scale parameter representing the overall variance. Given all neighbourhoods of area  $j$ ,  $N_j \equiv \{k: c_{jk} \neq 0\}$ , the model can be equivalently written as

$$\beta_{0j} \sim \text{Normal}(\gamma_{00} + \phi \sum_{k \in N_j} c_{jk}(\beta_{0k} - \gamma_{00}), vm_{jj}).$$

This model assumes that there is an association between the risks in areal unit  $j$  and the risks in neighbouring areal units, defined by the parameter  $\phi$  and the matrix  $C$ . The choice of the variance-covariance matrix  $v\mathbf{\Sigma}$  determines how the spatial association is expressed in the model. As noted above, the parameters  $\phi$ ,  $C$ , and  $M$  must be chosen so that  $\mathbf{\Sigma}$  is symmetric and positive-definite.  $\mathbf{\Sigma}$  is symmetric iff  $C_{jk} M_{kk} = C_{kj} M_{jj}$ . To make the variance  $vm_{jj} > 0$ ,  $m_{jj}$  must be  $> 0$ . Furthermore, to ensure that  $\mathbf{\Sigma}$  is positive-definite,  $\phi$  must lie between  $\phi_{\min}$  and  $\phi_{\max}$  where  $\phi_{\min}^{-1}$  and  $\phi_{\max}^{-1}$  are the smallest and largest eigenvalues of  $M^{-1/2} C M^{1/2}$ .

A large number of choices for  $M$ ,  $C$ , and  $\phi$  exist to meet these required conditions. The most commonly used model for spatial association is the intrinsic conditional autoregression model (Besag et al., 1991), which is obtained by choosing

$$\begin{aligned}
m_{jj} &= |N_j|^{-1}, \text{ the inverse of the number of neighbourhoods of the } j^{\text{th}} \text{ areal unit} \\
c_{jk} &= \begin{cases} |N_j|^{-1}, & j \in N_j \\ 0, & \text{otherwise} \end{cases} \\
\phi &= 1 = \phi_{\max}.
\end{aligned}$$

This is equivalent to  $\beta_{0j} \sim \text{Normal}(\bar{\beta}_{0j}, v_j)$ , where  $\bar{\beta}_{0j} = \frac{\sum_{j \neq k} w_{kj} \beta_{0k}}{\sum_{j \neq k} w_{kj}}$  and  $v_{0j} = \frac{v}{\sum_{j \neq k} w_{kj}}$ , where  $w_{kj}=1$  if unit  $j$  and  $k$  are adjacent, and  $w_{kj}=0$  otherwise. The variance-covariance matrix  $v\Sigma$  is singular at  $\emptyset_{\max}$ , leading to an improper distribution of  $\beta_{0j}$ . This specification can therefore only be used as a prior distribution for spatially distributed random effects, and not as a likelihood measure for data. It is often convenient to assume that such random effects have zero mean. The term  $\beta_{0j}$  is therefore partitioned into the random effects  $u_{0j}$  with zero mean, and a separate intercept term,  $\gamma_{00}$ , with a location invariant uniform prior (Besag and Kooperberg, 1995). The conditional variance for the areal unit  $j$  is therefore  $v / N_j$ , meaning that the variance decreases as number of neighbourhoods increases.

Another choice for a proper CAR variance matrix is from Stern and Cressie (1999), namely

$$\begin{aligned} m_{jj} &= E_j^{-1}, \text{ the inverse of the expected count in the } j^{\text{th}} \text{ areal unit} \\ c_{jk} &= \begin{cases} \left( \frac{E_k}{E_j} \right)^{1/2}, & k \in N_j \\ 0, & \text{otherwise} \end{cases} \\ \emptyset &\in (\emptyset_{\min}, \emptyset_{\max}), \end{aligned}$$

where  $E_j$  and  $E_k$  are expected values of occurrences in area  $j$  and  $k$  respectively.

Cressie and Chan (1989) used an alternative specification to incorporate a measure of the distance between two areal units, which can be specified as

$$\begin{aligned} m_{jj} &= n_j^{-1}, \text{ the inverse of the expected count in the } j^{\text{th}} \text{ areal unit} \\ c_{jk} &= \begin{cases} D(j) d_{jk}^{-p} \left( \frac{n_k}{n_j} \right)^{1/2}, & k \in N_j \\ 0, & \text{otherwise} \end{cases} \\ \emptyset &\in (\emptyset_{\min}, \emptyset_{\max}), \end{aligned}$$

where  $n_j$  and  $n_k$  are the population counts in areas  $j$  and  $k$  respectively. The parameter  $p$  is specified to be 0, 1, or 2 according to how fast it is desired that  $c_{jk}$  decreases with distance  $d_{jk}$ . The factor  $D(j)$  is a constant of proportionality defined as  $D(j) = (\min\{d_{jk} : j \in N_j\})^p$  so that  $c_{ij}$  are comparable across different values of  $j$ . The benefit of specifying weights in terms of population ratios is that the conditional variance for areal unit  $j$  is  $v/n_j$  so that the variance decreases as population size increases.

Since the spatial random process may not be perfectly structured, it is often unclear in practice whether to choose an unstructured random effect as specified in Equation (4.14), or a purely structured random effect as in Equation (4.23). An intermediate distribution that ranges from independence to local dependence, called a convolution prior model (Besag, 1989), has been proposed to deal with this situation. The model can then be defined as

Adverse birth outcomes  $\sim \text{binary}(p_{ij})$

Level 1 (personal):  $\text{logit}(p_{ij}) = \beta_{0j} + \beta_1 X_{ij}$

Level 2 (neighbourhood):  $\beta_{0j} = \gamma_{00} + u_{0j} + v_{0j}$ , (4.24)

where  $v_{0j}$ , as denoted in Equation (4.14), has a normal distribution with zero mean and variance  $\lambda^2$ , describing the unstructured heterogeneity of the relative risks. The term  $u_{0j}$  follows a CAR model with conditional variances  $vm_{jj}$ , representing local spatially structured variation. Thus, the variance of  $\beta_{0j}$  is  $\lambda^2 + vm_{jj}$ . This model takes into account the hierarchical nature of the risks of adverse birth outcomes and can incorporate the impacts of covariates at both the personal and neighbourhood levels. The spatially structured random effects are separated from the unstructured random effects. Thus, the spatial variation of the risks of adverse birth outcomes can be effectively modeled and identified.

It is difficult to fit this model using a marginal likelihood approach. Hence, a Bayesian framework is used to model these random effects. In this context, the Bayesian approach combines two types of information, namely the information provided by the covariate-adjusted adverse birth outcomes by the binomial likelihood  $[\mathbf{y}|\mathbf{r}]$ , and prior information on the relative risks specifying their variability among individuals and among areal units, summarised by their prior distribution  $[\mathbf{r}]$ . The prior for the area-specific random effects  $\beta_{0j}$  is specified by the above unstructured and spatially structured distributions. The prior for the parameter  $\beta_1$  is specified by a non-informative normal distribution with a large variance.

The Bayesian inference about the unknown relative risks  $\mathbf{r}$  is then given by the marginal posterior distribution  $[\mathbf{r}|\mathbf{y}] \propto [\mathbf{y}|\mathbf{r}] \times [\mathbf{r}]$ . Since the distribution of adverse birth outcomes  $y_i$  depends only on  $r_i$ , the  $y_i$  are conditionally independent given  $r_i$ . The likelihood function of  $\mathbf{r}$  for the data is then the product of  $n$  independent binomial distributions

$$[\mathbf{y}|\mathbf{r}] = \prod_{i=1}^n [y_i|r_i].$$

The prior distributions  $[\mathbf{r}]$  are parameterised by hyper-parameters  $\gamma$  and denoted  $[\mathbf{r}|\gamma]$ . For example, the hyper-parameters for a non-informative normal prior will be the mean and the variance. The joint posterior distribution of the parameters  $(\mathbf{r}, \gamma)$  is then  $[\mathbf{r}, \gamma|\mathbf{y}] \propto [\mathbf{y}|\mathbf{r}] \times [\mathbf{r}|\gamma] \times [\gamma]$ . The marginal posterior distribution is then

$$[\mathbf{r}|\mathbf{y}] = \int [\mathbf{r}, \gamma|\mathbf{y}] d\gamma \propto \prod_{i=1}^n [y_i|r_i] \times [\mathbf{r}|\gamma] \times [\gamma].$$

A point estimate of the parameters and the distribution of the spatially structured random effects is given by locating this posterior distribution and selecting the posterior mean  $E[\mathbf{r}|\mathbf{y}]$  or the posterior

median. However, it is generally impossible to evaluate these parameters directly through analytic or numerical integration. Instead, Markov Chain Monte Carlo (MCMC) methods are generally used to obtain the posterior distribution.

The MCMC methods use the general Metropolis algorithm to simulate a Markov chain. The equilibrium distribution of this chain is the desired distribution (Gilks et al. 1996). An adaptation of this algorithm is Gibbs sampling (Geman and Geman, 1984), which is particularly suitable when the joint posterior distribution is complicated but the full conditional distributions have simple forms.

Given the current values for other parameters, Gibbs sampling visits each parameter, including  $\beta_{0j}$ ,  $\beta_1$ , and hyper-parameters  $\gamma$  in turn, and simulates a new value from its full conditional distribution. For example, a new value of the parameter  $r_i$  is drawn from the full conditional distribution given the current values of other parameters ( $r_j', j \neq i$ ) and hyper-parameters  $\gamma'$

$$[r_i | \mathbf{y}, \mathbf{r}_j', j \neq i, \gamma'] \propto [y_i | r_i] \times [r_i | \mathbf{r}_j', j \neq i, \gamma'],$$

and a new value of  $\gamma$  is drawn from the full conditional distribution given the current values  $\mathbf{r}$

$$[\gamma | \mathbf{r}'] \propto [\mathbf{r}' | \gamma] \times [\gamma].$$

Theoretically, after a sufficiently long run of fitting this algorithm with sample data, the joint distribution of the sample values converges to the joint posterior distribution when the chain is irreducible, aperiodic and positive recurrent. Hence, the distribution of the sample values of  $\mathbf{r}$  converges to the marginal posterior distribution of  $[\mathbf{r} | \mathbf{y}]$  (Roberts, 1996; Tierney, 1996).

There is currently no a golden standard to diagnose and conclusively determine model convergence. However, certain practical guidelines can be taken for assessing convergence. It is recommended to use a combination of convergence diagnostics plus visual inspection and summary statistics to look for evidence of model stabilization. First, convergence can be monitored by visual examination of the trace plots or the history plots of the sample values versus iteration to look for evidence of stabilization. The variable is stabilized when sample values are randomly distributed above and below a stable mean value. If more than one chain is run simultaneously and all the chains appear to be overlapping with each other, it is reasonably confident that convergence has been achieved (David et al., 2003). Second, the autocorrelation graphs plots the autocorrelation function of the variable. It helps to determine the thinning number of a chain to get independent samples. The thinning number is the number to select every  $k^{th}$  iteration of each chain to contribute to the statistics being calculated. A high value of the autocorrelation coefficient (close to 1) indicates slow mixing and a large thinning number should be selected to ensure sample independence for statistical inference.

Third, some statistical diagnostics can also be found for convergence testing. A common

diagnostic used in WinBUGS is the Brooks, Gelman and Rubins (BGR) statistic, which is developed by Gelman and Rubins (1992) and modified later by Brooks and Gelman (1998). This diagnostic is appropriate for the analysis of two or more parallel chains, each starting at different over-dispersed initial values with respect to the target distribution. Convergence is assessed by comparing within- and between-chain variance over the second half of those chains to estimate the potential scale reduction factor, which might be reduced if the chains approach infinity. The diagnostic is defined as:

$$R = B / W$$

where, in WinBUGS, 80% credible interval is taken for the parameter of interest and intervals can then be calculated from the second half iteration of each chain or all chains.  $W$  denotes the within-chain variance calculated by the average width of empirical credible intervals across the chains.  $B$  is the between-chain variance calculated by pooling all samples from all chains together. If the ratio  $R$  greater than 1, the values are over-dispersed. As convergence is approached,  $R$  will tend to 1. While the BGR statistic is good at the analysis of two or more parallel chains, other diagnostics that are appropriate for the analysis of individual chains can also be found, such as the Geweke convergence diagnostic (Geweke, 1992) and the Heidelberger and Welch (1983) convergence diagnostic.

The choice of the hyperprior distribution of the parameters is another key issue since Bayesian models are sometimes sensitive to the selection of prior distributions. Sensitivity analyses should be carried out to investigate how much impact on posterior inference with reasonable modification of priors or parameters (Law and Haining, 2004). The selection of priors is in principle subjective. If certain prior knowledge can be obtained from, for example, experts or existing data, it can be formatively selected although the assumed relevance is still a subjective judgment. In a fully Bayesian approach to hierarchical modeling, prior information for the variances of parameters of interest usually cannot be obtained. It is necessary to select a non-informative prior distribution. Commonly used prior distributions in WinBUGS to model the precisions of parameter values include Gamma (0.001, 0.001), Gamma (0.5, 0.0005), and uniform distributions with reasonably wide ranges. By sensitivity analysis to a number of priors, assessments can be given on whether the current results are convincing to a broad spectrum of opinion.

After convergence is achieved, dependent sample values can be generated from the posterior distribution to approximate the distribution. The accuracy of resembling the posterior distribution is determined by the sample sizes selected after convergence. The question is how many iterations should be taken after convergence? A rule of thumb is that the Monte Carlo error for each parameter of interest should be less than about 5% of the sample standard deviation. Once sufficient samples are simulated, point estimates can be obtained from the simulated sample values and the posterior mean

can be approximated by the sample mean. The posterior median can also be represented by the sample median. Interval estimation is also available by computing Bayesian credible intervals from these samples. Several diagnostics have been proposed to assess convergence (Cowles and Carlin, 1996).

#### 4.3.1.iii Spatial scan statistics

Both of the above global and local indices are calculated based on predefined areal units and are greatly affected by the scale of spatial units being used. Calculating the spatial autocorrelation at different spatial scales or different divisions of spatial units may yield different, sometimes controversial, results. This is often called a modifiable areal unit problem (MAUP) (Openshaw and Taylor, 1979). The MAUP arises from the assumptions that the distance of individuals  $a$  and  $b$  equals 0 for all individuals in the same unit and that the distance of individuals  $b$  and  $c$  equals the distance of units  $i$  and  $j$  for all  $b$  in  $i$ ,  $c$  in  $j$ , and  $i \neq j$ . However, this is only true when all individuals in a given areal unit are located at the same place. The smaller the areal units, and the closer the boundaries between units are to real social and spatial barriers, the more reasonable this assumption is.

The clustering of adverse birth outcomes may occur at a certain spatial scale or at several spatial scales, caused by different levels of social and environmental determinants. For example, adverse birth outcomes may be found clustered at the city-wide level due to the impact of different health intervention policies among cities. They may also be clustered at a very small spatial scale, such as neighbourhoods or sub-neighbourhoods, due to the impact of adverse living environments. Thus, although having data aggregated at smaller areal units may reduce the impacts of the MAUP and be more likely to reveal the true spatial structure, spatial autocorrelation of low birth weight needs to be examined at different spatial scales.

Since spatial units may be artificially delineated based on different purposes, their divisions may not reflect the real situation that causes the clustering of low birth weight incidence. Thus, the analysis of spatial autocorrelation based on these spatial units may not yield satisfactory results. In this case, a spatial point process analysis may be invoked to detect clusters and assess the significance of them. The spatial scan statistic (Kulldorff, 1997) is commonly used in such cases to test whether a spatial point process is purely random or if any clusters can be detected. The scan statistic allows clusters of any shapes and sizes to be detected.

Either a Poisson or a Bernoulli model can be assumed for the underlying spatial point process. The aim is to test whether or not the point events are clustered in a subset  $Z$  within the study region  $R$ . For the Bernoulli model, each individual (or observation) within the zone  $Z \subset R$  has probability  $p$  of being a point (or being a success for an event), while the probability of success for individuals outside

the zone is  $q$ . The probability for any one individual is independent of all others. The null hypothesis is  $H_0: p=q$  and the alternative hypothesis is  $H_1: p>q$ . Under  $H_0$ , the distribution of the number of points is  $N(A) \sim \text{Binomial}(n_A, p)$  for all sets  $A$  in  $R$ . Under  $H_1$ , the point distribution is  $N(A) \sim \text{Binomial}(n_A, p)$  for sets  $A \subset Z$ , and  $N(A) \sim \text{Binomial}(n_A, q)$  for set  $A \subset Z^c$ , where  $Z^c$  is the complement of  $Z$  in  $R$ . For the Poisson model,  $N(A) \sim \text{Poisson}(pn_Z + qn_{Z^c})$ . The null hypothesis is  $H_0: p=q$ , and  $N(A) \sim \text{Poisson}(pn_Z)$  under this hypothesis. The alternative hypothesis is  $H_1: p>q$ .

The likelihood function for the Bernoulli model can be expressed as

$$L_Z = \begin{cases} L_{\hat{Z}} = (p)^{o_Z}(1-p)^{n_Z-o_Z}(q)^{o_G-o_Z}(1-q)^{(n_G-n_Z)-(o_G-o_Z)} & \text{if } p > q \\ L_0 = \left(\frac{o_G}{n_G}\right)^{o_G} \left(1 - \frac{o_G}{n_G}\right)^{n_G-o_G} & \text{otherwise,} \end{cases} \quad (4.25)$$

where  $o_Z$  and  $o_G$  are the observed occurrences of events in  $Z$  and  $G$  respectively,  $n_Z$  and  $n_G$  are the total number of observations in  $Z$  and  $G$ ,  $p = o_Z/n_Z$  and  $q = (o_G - o_Z)/(n_G - n_Z)$ . The most likely cluster can then be calculated accordingly. To make statistical inferences, the likelihood ratio,  $\lambda = L_{\hat{Z}}/L_0$  is used as the test statistic.

Similarly, for the Poisson model, the likelihood function is calculated as

$$L_Z = \begin{cases} L_{\hat{Z}} = \frac{e^{-pn_Z-q(n_G-n_Z)}}{o_Z!} p^{o_Z} q^{o_G-o_Z} \prod_i n_i & \text{if } p > q \\ L_0 = \frac{e^{-o_G}}{o_G!} \left(\frac{o_G}{N_G}\right)^{o_G} \prod_i n_i & \text{otherwise,} \end{cases} \quad (4.26)$$

where  $n_i$  is the number of observations in location  $i$ . The likelihood ratio is similarly defined as,  $\lambda = L_{\hat{Z}}/L_0$ , which can be used as the test statistic for statistical inference.

When the total number of points is small compared to the total number of observations, in other words, when the success rate is low, the Bernoulli model and the Poisson model closely approximate each other (Kulldorff, 1997). In this case, the choice of one model over the other does not make much difference. In other cases, if binary outcomes (such as personal adverse birth outcomes) are obtained, the Bernoulli model may be preferred. If count data (such as adverse birth outcome counts in each small area) are available, the Poisson model is a better choice.

The distribution of the test statistic does not have a closed analytical form. A Monte Carlo simulation is usually used after the value of the test statistic is calculated (Turnbull et al., 1990). Conditional on the known underlying observations,  $n_G$ , replications of the data set generated under the null hypothesis can be obtained. With a large number of such replications, such as 999 or 9999, the test can be given by the distribution of the simulated replications. If the value of the test statistic



for the real data set is among the upper 5 percent of all values, it can be said that the identified cluster is significant at the 95 percent confidence level.

In addition to the most likely cluster, various secondary clusters with high likelihood values may also be detected by the scan statistic. Some of these clusters are closely related to the most likely cluster with overlapping zones and a similar set of points. These secondary clusters provide little further information on the clusters within the study area. Some other clusters may be located in areas other than the location of the most likely cluster and do not overlap with it. These clusters provide additional information on potential clusters and are worth reporting along with the most likely cluster. A limitation of the scan statistic is that it is not easy to adjust for covariates in the test, especially for the Bernoulli model. While the identified clusters and cluster sizes can be mapped, the overall distribution of relative risks cannot be easily mapped using the results of the scan test.

The various local clustering analysis statistics provided above answer the questions of where incident cases are clustered in space and whether or not the identified clusters are statistically significant. However, as discussed above, no one statistic is without its limitations. For a specific case, such as the test of adverse birth outcome clusters and the spatial distribution of relative risks, a combination of different local spatial clustering analyses is necessary. Based on the advantages it offers, the convolution prior model denoted by Equation (4.24) and spatial scan statistics are used in Chapter 5 to assess local clustering of adverse birth outcomes.

#### **4.3.2 Spatial regression analysis**

The impact of space and place is not only evident on the spatial variation of adverse birth outcomes. It is also apparent in the associations of risk factors and outcomes. The impacts of some identified risk factors (both individual and contextual) on adverse birth outcomes may exhibit spatial non-stationarity. In other words, the impacts or the exposure-response associations are not fixed and they may vary from place to place. Several reasons may explain the appearance of some spatial variations of the associations (Fotheringham et al., 2002).

The first cause of spatial non-stationarity may be sampling variation. Even though there may be no underlying spatial non-stationarity in a data set, the parameter estimates obtained from models within each of the areal units will not be the same due to different samples of data being used in different areal units. Since this is a statistical artifact, the variation is not useful for explaining any of the underlying spatial processes, but it does need to be considered so that substantive causes of spatial non-stationarity can be distinguished.

The second possible cause is that some relationships are intrinsically different across space. The social and geographical theories discussed earlier in Chapter 3, such as structuration theory and social

ecological theory, may help to explain the spatial processes that cause non-stationarity. Different socio-economic, political, structural, behavioural and environmental determinants may vary spatially. Moreover, personal responses to these determinants may also be different due to different personal characteristics. This makes it hard to separate the effects of social determinants in explaining health variation, and therefore a systematic view is needed. Within this framework, the identification of local variations of personal risks would be a useful precursor for qualitative or intensive studies to explain why such spatial differences occur. It may also be useful for a black-box intervention for area-based local health promotion.

The third possible cause might be a gross misspecification of reality, with one or more relevant variables being omitted from the causal model. Counter to reasoning for the second cause, this explanation holds a positivist view. It believes that a global statement can be made to explain spatial variations but that the structure of a current model may not be sufficiently well formed to allow all the determinants to be properly considered in making the global statement. Mapping local variation is useful for understanding the misspecification and finding out what additional attributes should be included in order to improve the model. In this sense, the goal of local analysis is to produce a global model that exhibits no significant spatial non-stationarity. Alternatively, if it is impossible to reduce the misspecification problem due to the unavailability of information on some variables, local modeling can serve to allow these otherwise omitted effects to be included.

Various spatial modeling techniques have been proposed to measure the spatial non-stationarity of exposure-response associations. These are now discussed.

#### 4.3.2.i The geographically weighted regression method

Geographically weighted regression (GWR) is commonly used to model spatial processes. It is based on two basic geographical assumptions. One is that spatial variation changes continuously from one location to another. The other is that spatial processes determine that observations are not independent of each other in space. Hence, observations in close spatial proximity to one other are more than likely correlated.

If an OLS model is defined as in Equation (4.1), a corresponding GWR model can be written as (Fotheringham et al., 2002):

$$Y_i = \beta_0(u_i, v_i) + \beta_1(u_i, v_i)X_i + e_i, \quad (4.27)$$

where  $(u_i, v_i)$  denotes the coordinates of the  $i^{th}$  point in space and  $\beta(u_i, v_i)$  is a realization of the continuous function  $\beta(u, v)$  at point  $i$ . It therefore allows there to be a continuous surface of parameter values. Equation (4.1) can be seen as a special case of Equation (4.27) in which the parameters are assumed to be the same across space. In essence, the equation measures the relationships around each

location  $i$ , which can be any location in space, not necessarily the location of an observation.

As noted above, the calculation of Equation (4.27) assumes that observations near to location  $i$  have more influence in the estimation of  $\beta(u_i, v_i)$  than do observations located farther from  $i$ . An observation is therefore weighted in accordance with its proximity to the location  $i$ . Observations close to  $i$  are weighted more than observations farther away.

In Equation (4.1),  $\beta_0$  and  $\beta_1$  are estimated by:

$$\beta_0 = \Sigma y_i / n \quad \beta_1 = \Sigma x_i y_i / \Sigma x_i^2.$$

Similarly,  $\beta_0(u_i, v_i)$  and  $\beta_1(u_i, v_i)$  in Equation (4.27) can be calculated as:

$$\beta_0(i) = \Sigma W(u_i, v_i) y_i / \Sigma W(u_i, v_i) \quad \beta_1(i) = \Sigma W(u_i, v_i) x_i y_i / \Sigma W(u_i, v_i) x_i^2,$$

where  $W(u_i, v_i)$  denotes the geographical weighting of each of the  $n$  observed data for regression point  $i$ . Thus, a GWR estimates the relationship for each location  $i$  using the OLS method on the observations that are around  $i$ . The values of observations close to  $i$  are weighted according to their proximity to  $i$  before they are incorporated into the calculation.

$W(u_i, v_i)$  can take various forms depending on the conditions in the study area and the nature of the data being used. If  $W(u_i, v_i)$  equals 1, the GWR model reduces to the OLS model. If it equals 1 within a certain distance of  $i$ , and 0 otherwise, the GWR is similar to a moving window approach, in which estimation in location  $i$  is made according to observations that are within a certain distance, and these observations are not weighted. It can also be a fixed kernel in which a distance-decay function is used to estimate the weighting. As sample data are not equally distributed over space, the problem with the fixed kernel is that in areas where data are dense, the fixed kernels are larger than they need to be and hence the estimates may be biased. Spatially varying kernels can be used to overcome this limitation.

A limitation of the GWR model is that it models associations at only one spatial level. In modeling disease risks, the multilevel nature of risks cannot be effectively modeled by GWR. To be able to take into account the hierarchical impacts of health risks, multilevel spatial modeling techniques can be used to model the spatial variations of exposure-response associations at different levels.

#### 4.3.2.ii Multilevel spatial models

Using the idea of the convolution prior model in Equation (4.24), a multilevel spatial model to measure the spatial distribution of personal risks of adverse birth outcomes can be specified as

Adverse birth outcomes  $\sim \text{binary}(p_{ij})$

Level 1 (personal):  $\text{logit}(p_{ij}) = \beta_{0j} + \beta_{1j} X_{ij}$

$$\begin{aligned}\text{Level 2 (neighbourhood): } \beta_{0j} &= \gamma_{00} + u_{0j} + v_{0j} \\ \beta_{1j} &= \gamma_{10} + u_{1j} + v_{1j}.\end{aligned}\tag{4.28}$$

Instead of modeling the parameters of personal risks as fixed effects, this model specifies that these parameters may also vary spatially. The spatial variations of the parameters,  $\beta_{1j}$ , are modeled as neighbourhood-level unstructured and spatially structured random effects,  $v_{1j}$  and  $u_{1j}$ . The spatially structured random effect  $u_{1j}$  can be similarly constructed as following one of the CAR models described in the last section. This model suggests that the spatial variation of personal-level associations is due to both random sampling variation and gross misspecification of reality. The spatially structured random effect is due to unmodeled relevant neighbourhood-level risks, which are also supposed to be spatially autocorrelated. Due to the complexity of the social determinants of health and the interactions between them, it is sometimes hard to separate the effects of social determinants and incorporate them effectively into the model even though some of them may be identified to have causal inferences (Lawson, 1999).

This model is flexible in the sense that it is able to incorporate any identified risk factors and model the spatial impacts of unidentified risks on both the health outcomes and the identified risk factors. For example, if certain neighbourhood risks,  $W_j$ , can be identified as having fixed impacts on adverse birth outcomes, they can be incorporated into the model as

$$\begin{aligned}\text{Adverse birth outcomes} &\sim \text{binary}(p_{ij}) \\ \text{Level 1 (personal): } \logit(p_{ij}) &= \beta_{0j} + \beta_{1j}X_{ij} \\ \text{Level 2 (neighbourhood): } \beta_{0j} &= \gamma_{00} + \gamma_{01} W_j + u_{0j} + v_{0j} \\ \beta_{1j} &= \gamma_{10} + u_{1j} + v_{1j}.\end{aligned}\tag{4.29}$$

If it is believed that the impacts of neighbourhood risk  $W_j$  are also different from city to city, a three level model may be constructed as:

$$\begin{aligned}\text{Adverse birth outcomes} &\sim \text{binary}(p_{ij}) \\ \text{Level 1 (personal): } \logit(p_{ij}) &= \beta_{0j} + \beta_{1j}X_{ij} \\ \text{Level 2 (neighbourhood): } \beta_{0j} &= \gamma_{00} + \gamma_{01k} W_j + u_{0j} + v_{0j} \\ \beta_{1j} &= \gamma_{10} + u_{1j} + v_{1j} \\ \text{Level 3 (municipality): } \gamma_{01k} &= \alpha_{011} + u_{01k} + v_{01k}.\end{aligned}\tag{4.30}$$

Due to their complexity, these models can only be fitted using Bayesian frameworks. The MCMC methods described in last section can be used to obtain the posterior distribution for statistical inference. Instead of a full Bayesian framework, these models may also be fitted with empirical Bayesian methods based on certain assumptions of the parameters. These models and their

assumptions are discussed in the next section.

#### 4.3.2.iii Empirical Bayesian estimation for local variance

In Equation (4.28), the full Bayesian methods assign informative priors to the parameters  $\beta_{0j}$  and  $\beta_{1j}$ , and estimate them using the full conditional posterior distribution. The Empirical Bayesian (Armitage and Colton, 1998) idea consists of approximating the parameters by suitable estimates of the prior. For example, if  $\beta_{0j}$  and  $\beta_{1j}$  are estimated for a particular neighbourhood, an ordinary least square (OLS) regression (or generalized linear (GL) regression model for binary outcomes) can be used for individuals in that specific area and take the parameters of  $\beta_{0j}$  and  $\beta_{1j}$  from the local OLS or GL results as the estimations of level-1 predictors in that area. However, this would be an effective estimate for the particular area only if Y were measured with no error, since there is less data in local areas than the whole study region and there is a possibility that the local variation is only caused by chance. If the results from a local OLS or GL estimation are not thought to be accurate, the estimation of  $\beta_0$  and  $\beta_1$  from the global OLS or logistic model that uses all of the data will be a good estimation. The question here is which of the above two estimates can be trusted. The Empirical Bayes model makes a compromise between these two estimations, which means that the less reliable the local OLS or GL measure of Y in a particular area is, the more the grand estimation across all areas can be trusted as an accurate estimate. The empirical Bayes estimates of  $\beta_{0j}$  and  $\beta_{1j}$  can then be defined as

$$\beta_{0j} = \lambda_j \beta_{0j}^{LM} + (1 - \lambda_j) \beta_0^{GM} \quad \beta_{1j} = \lambda_j \beta_{1j}^{LM} + (1 - \lambda_j) \beta_1^{GM}, \quad (4.31)$$

where  $\beta_{0j}^{LM}$  and  $\beta_{1j}^{LM}$  are the parameters of the local regression model  $\beta_0^{GM}$  and  $\beta_1^{GM}$  are the parameters of the global regression model, and  $\lambda_j$  is the reliability coefficient in areal unit  $j$ , which is calculated as

$$\lambda_j = \frac{\sigma_{v_0}^2}{\left(\sigma_{v_0}^2 + \frac{\sigma_r^2}{n_j}\right)},$$

where  $\sigma_{v_0}^2$  is the variance of neighbourhood-level residue,  $\sigma_r^2$  is the individual variance which is  $\pi^2/3$  for the logistic model, and  $n_j$  is the number of individuals in neighbourhood  $j$ . This reliability coefficient indicates that the major determinant of high reliability for a particular group is the number of individuals in that group ( $n_j$ ). This is to say that a neighbourhood with a relatively large sample size will contribute more information to the model than a group that only has a small number of individuals.

If the area variation is high and it explains most of the total variation, the reliability is high (close to 1). This means that there is obvious spatial non-stationarity and the local estimation is more

reliable. The Bayesian estimates will then be very close to the local OLS or GL estimates in the individual area. On the other hand, if the reliability is low, the Empirical Bayes estimate will approach the value from the grand OLS or GL estimation since most of the local variation may due to the global effects. The Empirical Bayes estimate will always lie between the local mean and the global mean. Thus, it represents a balance between the information obtained from a specific area and from the entire data set.

Other spatial regression models are also proposed by researchers. The spatial expansion method (Casetti, 1997) uses directly the coordinates of the samples in model building to measure spatial impacts. For example, the OLS model in Equation (4.1) can be modified to allow the parameters  $\beta_0$  and  $\beta_1$  to vary. Specifically,

$$\begin{aligned}\beta_{0i} &= \gamma_0 + \gamma_{01}x_i + \gamma_{02}y_i \\ \beta_{1i} &= \gamma_1 + \gamma_{11}x_i + \gamma_{12}y_i,\end{aligned}\tag{4.32}$$

where  $x_i$  and  $y_i$  represent the spatial coordinates of location  $i$ . This model allows expansion of the global parameters to associate linearly with spatial coordinates. More complex non-linear expansions can also be accommodated. Other spatial autoregression models, such as the spatial lag model (Anselin, 2002) and the spatial error model (Haining, 2003), use global parameters to measure the spatial dependency. While these models are useful in controlling for spatial impacts in order to make better inferences for point estimates of the covariates, they cannot measure the spatial variation of the parameters.

Given its flexibility in modeling spatial dependence, the multilevel spatial models discussed above are used in this thesis to assess the spatial variation of associations of adverse birth outcomes. By adjusting model constructions as described in Equations (4.28) through (4.30), the interrelation of social and spatial risks can be effectively modeled and the distribution of spatial variation of adverse birth outcome risks can be mapped to provide evidence for targeting health intervention strategies.

Overall, the local spatial analysis statistics discussed in this section provide an effective means to test the last hypothesis stated in Chapter 3 on whether or not the social and environment impacts of adverse birth outcomes are spatially dependent and whether they should be dealt with globally or locally when addressing their impacts and making decisions on health interventions.

#### 4.4 Preliminary analyses

In order for the proposed spatial analysis to be conducted, preliminary analyses were needed for data preparation. Individual data needed to be geo-referenced so that the locational information could be used in subsequent statistical modeling. Further, data obtained from various secondary sources

were not always suitable for direct use. Some needed to be manipulated so that effective personal and neighbourhood-level indicators could be constructed for statistical modeling. Specifically, factor analysis methods were used to obtain independent personal-level risk factors, and spatial interpolation methods were employed to obtain neighbourhood-level health-related indicators from health survey data. The following sections describe each of these methods respectively.

#### **4.4.1 Geocoding**

Depending on the nature of the spatial information obtained from the data, different geocoding methods can be used. The two typical geocoding methods are geocoding by street address and geocoding by postal codes. These methods provide the means to convert spatially related information to spatial locations.

##### **4.4.1.i Geocoding by street address**

Geocoding based on a street address is currently one of the most accurate geocoding methods for identifying residential households and their occupants. It involves matching addresses in a table to the street names and address ranges in a digital street network file. First, address matching is performed based on a street address locator. It matches the street name in both the table to be geocoded and a reference table of the street network, which is accompanied by a digital map of the street network arcs (segments). Once the street name is matched, street segments with the same street name in the reference map are extracted. The particular street address number is compared with all address ranges for the extracted street segments in order to identify the street segment where the address is located, on the even or odd numbered side (left or right) of the street. Second, based on the range of street numbers for the identified street segment and the coordinates of this segment's endpoints obtained in the reference map, the coordinates of the address to be geocoded can be interpolated based on a linear interpolation method. The final spatial coordinates of the street address can then be determined based on the interpolated result. However, despite the fact that this is the most accurate methods of geocoding household addresses, the geocoded spatial locations may not be the true locations of the addresses. Street address geocoding produces only an approximation of the true location, although the locational error may actually be quite small.

The address locator is the main tool for address matching in the ArcGIS software package. An address locator is created based on a specific locator style and a reference street network. In this case, the "US Streets with Zones" and "US Streets" address locator styles were appropriate since the obtained addresses for subjects examined in this thesis are within Ontario and follow North American street naming conventions. Once created, an address locator contains the geocoding properties and

parameters, address attributes in the reference data, and the queries for performing a geocoding search. It can then be used for the above mentioned address matching process. If geocoding is for addresses within one city, the “US Streets” address locator style is sufficient since no two streets have the same street names in one city. However, if geocoding involves larger areas which contain several cities, the “US Streets with Zones” locator is necessary since common street names, such as Main Street, can be found in almost every city. Using the “US Streets with Zones” locator, only the records that match both the street addresses and the zones (municipalities or three-digit postal codes) are geocoded. This avoids the problem of misallocating records from one city to the other.

#### 4.4.1.ii Geocoding by postal code

Geocoding by postal code is another commonly used geocoding method. Although it is less accurate than the geocoding by the street addresses approach, postal codes are more readily available than street addresses, especially in survey data. Due to security concerns, street addresses of disease incidences may not be released even if they are collected. Hence, a geocoding by postal code method can be used in this case to obtain spatial information. Any database that includes postal code information can be geocoded based on a file that consists of postal code centroid points with geographic coordinates attached to them. In Canada, the Postal Code Conversion Files (PCCF and PCCF+) produced and maintained by Statistics Canada also provide a correspondence between the six-character postal code and Statistics Canada's standard geographic areas for which census data are produced. Through the link between postal codes and standard geographic areas, the PCCF permits the integration of spatial data from various sources.

Using the associated postal code, the spatial reference of a sample can simply be obtained by assigning a spatial point within its corresponding postal code area as the occurrence location, such as the centroid of the postal code area. However, if the purpose of geocoding is to identify in which census dissemination areas the sample events have occurred, and a postal code area overlaps with several census dissemination areas, it may be problematic to assign the spatial locations of the sample to the centroids of their corresponding postal code areas. All samples that have the same postal code will be assigned to the same spatial point and will therefore be assigned to one census dissemination area, which may generate non-existent peak values in that census dissemination area. The Postal Code Conversion Files provide a way to avoid this problem. Using PCCF+ files, if one postal code area overlaps with several census dissemination areas, samples that have the same postal code will be randomly assigned to a spatial point within one of these areas in proportion to the population of these areas. While this avoids samples with the same postal code being geocoded to one location, it is still prone to producing locational error.



Due to data quality and data collection issues (for example, not every subject may be willing to provide address information), one single geocoding method may not be sufficient to obtain sufficiently valid spatial referencing for all obtained data. Hence, a combination of different geocoding methods may be necessary to achieve a desired percentage of geocoded data. In this thesis, both the geocoding by street address and geocoding by postal code methods were used to obtain spatial locations of individual live births within the study area.

#### **4.4.2 Spatial interpolation**

The literature discussed in Chapter 2 revealed that a majority of neighbourhood or small-area analyses focus on the impacts of socio-economic status, ethnicity, demography, and physical living conditions and their consequences on adverse birth outcomes. This is largely due to the ready and broad-based availability of census data, which can be aggregated at different spatial levels to obtain the indicators mentioned above. However, other important aspects of social determinants of adverse birth outcomes, such as psycho-social risk factors, health service provision, food supply, walkable and liveable communities, urban design features, and average community health conditions, have not been fully investigated at the neighbourhood level due to the difficulties in obtaining corresponding indicators. Effective neighbourhood-level potential risk factors need to be identified and corresponding data need to be collected or derived at different spatial scales to make the analyses possible.

The Canadian Community Health Survey (CCHS) data provide useful information for constructing neighbourhood-level variables. These variables cover a wide range of health-related risks and may have potential uses for the analyses of social and environmental impacts on population health. However, since the primary purposes of the data are for analyses at the relatively larger health region level, they cannot be utilized directly in secondary data analyses due to sampling issues, especially when indicators are required at smaller areal units than the designed uses. CCHS surveys are designed and collected for community health research at the health region level, which is much larger than the neighbourhood level required for small-area analysis. Within each health region, samples are not randomly distributed among smaller areal units, such as census dissemination areas (CDA). Sample sizes are much larger in urban areas than in rural areas. At each survey cycle, a certain number of CDAs in rural areas do not contain even one sample point. Thus, it is not feasible to aggregate directly personal-level variables to obtain small area level indicators. These data therefore had to be effectively transformed before they could be used to explore the relationships between adverse birth outcomes and neighbourhoods in this thesis.

As discussed in Chapter 3, neighbourhoods represent spatial areas defined by people's places of

residence and the surrounding environments where people live and interact with each other. Influenced by personal choices or external forces, people with common socio-economic status, ethnicity or demographic characteristics tend to live close to each other within the same neighbourhood or in nearby neighbourhoods. They consider certain surrounding areas as comprising their neighbourhood or their community, within which community-level organizations may exist and common characteristics may be found that affect people's behaviours or beliefs. Neighbourhoods act as containers or constraints that limit local residents' access to service facilities, resources, education and job opportunities, constrain their interactions, and even the selection of marriage partners (Overall and Nichols, 2001), and affect and regulate their behaviours through common norms. Due to these impacts, homogeneities may be found within neighbourhoods or surrounding areas, in terms of living conditions, socio-economic status, health-related behaviours, psycho-social status, and general health.

Neighbourhood characteristics (or collectives) are often associated with personal-level characteristics (or members), but are not necessarily based on personal characteristics (Luke, 2004). Some collective characteristics can be obtained by aggregating information, such as the average income of an area, from individual members. Some characteristics are based on the relationships between members, such as the social network of a neighbourhood. Others are the characteristics of the collective itself, such as health care provisions or policies. However, the collective properties at the neighbourhood level are sometimes difficult to obtain, or at least to obtain directly. Samples for most health surveys are person-based, rather than collectivity-based. Some properties are multidimensional and hard to measure, such as social capital, which include aspects of social interactions, social network, trust, and mutual benefits. Some other properties may be affected by spatial and social distances and therefore may be hard to divide among neighbourhoods.

The purpose of this analysis is to obtain health-related neighbourhood properties through the aggregation of personal-level characteristics in surveys so that common characteristics may be found to represent collective properties. For example, social capital may be represented by local residents' average sense of belonging to local communities and the distribution of health provisions may be represented by the average unmet health care needs within each neighbourhood. Hence, based on the above discussion, it is reasonable to assume that individuals who live close to each other have similar socio-economic status, health-related behaviours, psycho-social stress or depressions, and health outcomes. This reality provides the theoretical foundation for using spatial interpolation methods to derive neighbourhood-level risk variables.

Spatial interpolation is used to predict values for specified spatial locations when only a limited number of sample data points can be obtained at nearby locations. What makes spatial interpolation a

viable option is that spatially distributed objects are often spatially correlated. In other words, things that are close together tend to have similar characteristics. This is consistent with the above mentioned theoretical assumptions on neighbourhoods. For instance, with spatial dependency, if people living on one side of the street are poor, it can be predicted with high confidence that people living on the other side of the street are poor as well. It will, however, be less certain to predict that people who live across the town are also poor given the large distance. This assumption or precondition of spatial dependency is the basis of many spatial interpolation methods. These methods can be grouped into deterministic and stochastic methods.

Deterministic methods interpolate values based on either the extent of similarity, such as the Inverse Distance Weighting (IDW) method, or the degree of smoothing as in the use of Radial Basis Functions, such as splines and multiquadric functions. The methods can be either global or local, and can be exact (such as IDW and spline) or inexact (such as polynomial interpolators). On the other hand, stochastic interpolation techniques, such as Kriging (Krige, 1951), quantify the spatial autocorrelation among sample points based on the entire dataset within the study area, and account for the spatial configuration of the sample points around the prediction locations for interpolation. Although Kriging is sometimes described as a local interpolation method since technically more weights are given to neighbouring observations, stochastic interpolation methods are in nature global interpolation techniques (Stein, 1999). The two most commonly adapted interpolators, namely IDW and Kriging, are now considered in more detail.

A widely used deterministic interpolation method is IDW. This is a local exact interpolator that interpolates values based only on the surrounding measured values of the interpolating location and functions of the inverse distances between the interpolating location and locations of the surrounding sample. Shepard's (1968) IDW interpolation method seeks to find an interpolated value  $z$  at a given point  $x$  based on samples  $z_i = z(x_i)$  for  $i = 0, 1, \dots, n$ . This method can be written as:

$$z(x) = \sum_{i=0}^n \frac{w_i(x)}{\sum_{i=0}^n w_i(x)} z_i, \quad (4.33)$$

where  $w_i(x) = \frac{1}{d(x, x_i)^p}$  is the weighting function denoted by the inverse Euclidean distance between the interpolating point  $x$  and the neighbouring data point  $x_i$ , raised to the power  $p$ .

Modifications to the weighting function for interpolation are also suggested. One modification is the Liszka's method (Liszka, 1984), which is formed by  $w_i(x) = \frac{1}{(d(x, x_i) + \delta^2)^p}$ , where  $\delta^2$  is a constant responsible for the measurement error to control the smoothness of interpolation. The interpolated surface is smoother along with the increase of  $\delta^2$  to obtain optimal approximation. The other

modification is the Lukaszuk-Karmowski metric (Lukaszuk, 2004), which defines  $w_i(x) = \frac{1}{D_{RD}(x, x_i)^p}$ , where  $D_{RD}(x, x_i)$  is the probability metric of random values  $x$  and  $x_i$  assuming they have uniform and Dirac delta distribution. This approach is physically based, allowing the actual distribution of the sample point around its location to be considered.

The IDW interpolation method depends on the selection of the power value,  $p$ , and the neighbourhood search strategy. It is an exact interpolator, where the interpolated surface goes through the sample points and maximum/minimum values in the surface only occur at sample point locations (Henley, 1981). IDW assumes that the spatial process at the interpolating locations is being driven by the local variation only, which can be captured through the surrounding sample points. The output values therefore are sensitive to local spatial clustering and the presence of outliers.

Stochastic methods, such as Kriging, interpolate values not only based on the surrounding data values, but also on the overall autocorrelation calculated by applying statistical models to all the known data points. Because of this, stochastic methods not only have the capability of producing a prediction surface, but they also provide some measure of the certainty or accuracy of the predictions. The Kriging method can be constructed as:

$$\hat{z}(x) = \sum_{i=1}^n w_i(x) z(x_i), \quad (4.34)$$

where the interpolated value  $z$  at point  $x$  is the weighted sum of the neighbouring observed values  $z_i = z(x_i)$  with weights  $w_i(x)$ ,  $i=1, \dots, n$ , chosen such that the Kriging variance is minimized, subject to the unbiasedness condition:

$$E[\hat{z}(x) - z(x)] = \sum_{i=1}^n w_i(x) \mu(x_i) - \mu(x) = 0,$$

where  $\mu(x) = E[z(x)]$  is the expected value of  $z(x)$ . This means that the interpolated value at a given location is the sum of two components, namely an unknown underlying surface defined by  $\mu(x)$ , plus some additional unaccounted noise.

Based on different assumptions on  $\mu(x)$  and the unbiasedness condition for calculating the weights, different types of Kriging can be used, including Simple Kriging, Ordinary Kriging, Universal Kriging, IRFk-Kriging, Indicator Kriging, Disjunctive Kriging, Lognormal Kriging, and co-Kriging.

The most commonly used type of Kriging is ordinary Kriging. It assumes a wide sense stationary process, in that the overall mean and autocorrelation remain relatively stable over time change. A constant but unknown mean,  $\mu(x) = \mu$ , is assumed to exist over the study region. Similar to the concepts of IDW interpolation, the weights,  $w_i(x)$ , decline as distances between the being-

estimated location and the data point locations increase. The difference is that, instead of using the local inverse distance function, the Kriging method uses the globally calculated semi-variogram (Wackernagel, 2003),  $\gamma(x,y) = E[(z(x)-z(y))^2]$ , to calculate the weights.

The semi-variogram is a function describing the degree of spatial dependence of a variable  $z(x)$ . It is usually represented as a graph that models the autocorrelation among sample locations to identify how values are varied with distance of separation. If there are enough observations and there is no directional effect, a semi-variogram can be empirically estimated by a semi-variogram cloud,  $\gamma(h)$ , which is defined as

$$\gamma(h) = \frac{1}{2} \frac{1}{n_h} \sum_{i=1}^{n_h} (z(x_i + h) - z(x_i))^2$$

where  $h$  is lag distance,  $n_h$  is the number of paired observations at the distance  $h$ , and  $z$  is the observed value at a particular location. The semivariance at lag distance  $h$ ,  $\gamma(h)$ , is half the variance of the differences  $z(x_i + h) - z(x_i)$ , which is equivalent to the whole variance of  $z$ -values at distance  $h$  (Bachmaier and Backes, 2008). If directional effect exists, the selection of paired observations to construct the semi-variogram is determined not only by their Euclidian distances but also by their spatial directions. The semi-variogram can then be plotted against a number of lags. To ensure validity for Kriging, the  $\gamma(h)$  are often approximated by model functions, such as exponential, spherical, and Gaussian functions.

Figure 4.4 shows a typical fitted semi-variogram, which is determined by the nugget, range, sill and the selection of the fitting functions. The horizontal axis represents distance between pairs of samples. The value of the semi-variogram at distance  $h$  is plotted along the vertical axis. The semivariogram might be non continuous at the origin due to random or measurement errors. The height of the jump at the origin is referred to as nugget or nugget effect. As distance increases, the

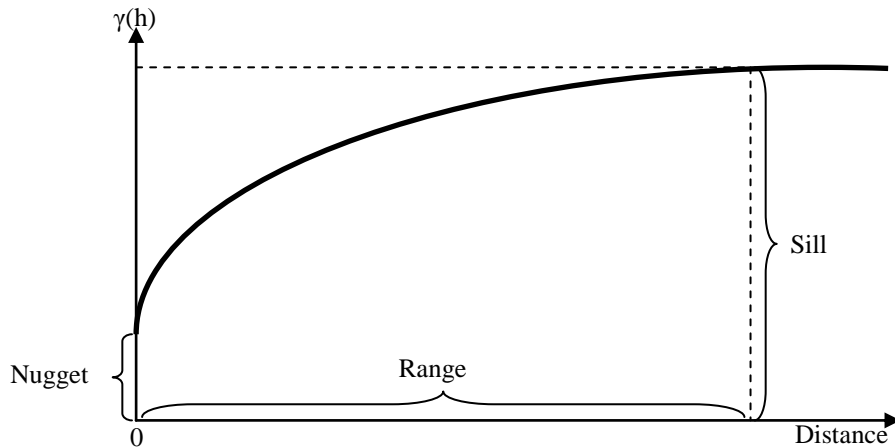


Figure 4.4: A typical fitted semi-variogram model

semi-variogram values increase until reaching the sill, which is the upper limit of the variogram when lag distances tend to infinity. Based on semi-variogram values at the lags, different fitting functions and parameters can be selected to obtain the best fit. Some of the functions do have a sill such as the spherical and Gaussian functions, while some others do not have a sill, such as the exponential function. For functions with a fixed sill, the range is defined as the distance at which the sill is first reached. For models with an asymptotic sill, the range is usually taken as the distance when the semi-variogram value first reaches 95% of the sill.

An appropriately fitted semi-variogram should be able to reveal the real scale-dependent spatial correlation. In addition to function and parameter selections, the choice of lag sizes and number of lags also affects the fit of the selected function to the empirical semi-variogram. Excessively large lag sizes may mask short-range autocorrelation while excessively small lag sizes may be associated with small sample sizes within each lag to achieve representative averages of  $\gamma(h)$ . The number of lags should also be selected carefully so that the range of the semi-variogram function is less than the maximum estimating distance obtained by multiplying the lag size by the number of lags. If the range of the fitted semi-variogram model is small relative to the extent of the empirical semi-variogram, the lag size or number of lags can be reasonably decreased to reduce estimating time without sacrificing general accuracy.

Once the semi-variogram model is fit, it can then be used to calculate the weights,  $w_i(x)$ , for Kriging. Since Kriging is estimated using neighbouring points of the predicting location, the neighbourhood structure or number of neighbourhood points also has potential impacts on the Kriging results. Ordinary Kriging assumes that the study region has a wide sense stationary process with a constant mean. This assumption is generally consistent with current regional-level health patterns in developed countries, such as Canada, where, due to the already advanced level of medical care and the fact that chronic diseases are the leading cause of mortality and morbidity, the overall health status and life expectancies remain relatively stable over time. Thus, neighbourhood differences in health-related status may be seen as spatially autocorrelated random effects above or below a constant average health status resulted by local influences. Given these factors, it is appropriate to use ordinary Kriging for interpolating the CCHS variables examined in this thesis.

If spatial autocorrelation exists, interpolation methods are suited for obtaining and deriving neighbourhood characteristics. The problem is then which method to use and what parameters are appropriate. Specifically, if the spatial autocorrelation is only a local process, IDW may be more appropriate than Kriging. Otherwise, if the spatial autocorrelation is universally applicable as a function of distance, Kriging should be applied.

Using a cross validation method, which takes a data point out of the fitting process and then

predicts its value and compares the prediction to its actual value, the accuracy and unbiasedness of different interpolation methods or different selection of parameters of the same method can be compared. In particular, the root-mean-square prediction error can be compared between different models as a way of choosing one model over another or adjusting parameter values. Since neighbourhood-level socio-economic status can be easily obtained from census data, it is also feasible to compare directly the interpolated CCHS socio-economic status variables, such as household income, with the corresponding variables from the census to demonstrate the accuracy of the spatial interpolation using CCHS data.

The two interpolation methods, IDW and Kriging, were compared for this thesis and appropriate procedures of interpolation were identified to obtain neighbourhood-level health related variables as explanatory variables in the statistical modeling.

#### **4.4.3 Factor analysis**

As identified in Chapter 2, various personal-level risks are associated with preterm births and LBW. Some of these risks are highly correlated, constraining the ability to create an integrated model to depict the overall impacts of personal-level risks on adverse birth outcomes. Due to collinearity issues, some of the personal-level variables need to be eliminated from the integrated regression model to obtain a fit with more stable coefficients. This may reduce some of the generalizability in interpreting the impacts of these variables on adverse birth outcomes.

To eliminate the difficulties caused by the many correlated dimensions of personal-level risks, factor analysis can be used to construct orthogonal composite personal-level risk factors. The benefits of a factor analysis include:

- identifying unique aspects of personal-level determinants of adverse birth outcomes so that different pathways maybe easily identified in later mediational analysis;
- handling collinearity issues effectively without losing generalizability; and
- reducing the dimensions of the spaces of personal-level impacts for model construction and calculation.

Factor analysis is a generic term for a family of techniques that reduce a set of observable variables to a small number of latent factors (Bartholomew, 1987). It was developed primarily for analyzing relationships among a number of measurable entities (such as survey variables). Factor analysis is based on the assumption that there exist unobserved latent variables, called factors, which account for the correlations among observed variables. If these latent variables are held constant, the partial correlations among observed variables become zero. In other words, the latent factors can determine the values of the observed variables. As discussed in Chapter 2, many different aspects of

personal risks of adverse birth outcomes can be indentified including medical, socio-economic, psycho-social, genetic, and behavioural risks. Summarizing the identified personal-level risk factors in terms of latent factors helps to improve construct validity of the multilevel models examined in this thesis.

In a factor analysis model, each observed variable can be expressed as a weighted composite of a set of latent variables (Khattree and Naik, 2000) as follows

$$x_i = \mu_i + l_{i1}f_1 + l_{i2}f_2 + \dots l_{in}f_n + e_i, \quad (4.35)$$

where  $x_i$  is the  $i^{th}$  observed variable,  $\mu_i$  is a constant representing the mean of  $x_i$ ,  $f_1$  through  $f_n$  are the latent variables or common factors,  $l_{i1}$  through  $l_{in}$  are the corresponding factor loadings, and  $e_i$  is the residual of  $x_i$  on the factors or specific factors. Given the assumption that  $e_i$  is uncorrelated across the observed variables, the correlations among the observed variables are therefore accounted for by the common factors.

Distinct factors are sought such that the correlations among the components of  $x$  are completely accounted by these factors. This amounts to saying that  $E(f)=0$ ,  $E(e)=0$ ,  $cov(f,e)=0$ ,  $D(f)=\Delta$ ,  $D(e)=\Psi =diag(\Psi_1, \dots, \Psi_I)$ , covariance of  $f$  and  $e$ ,  $cov(f,e)$ , is 0, the variance-covariance matrix  $D(f)$  is positive definite, and  $D(e)$  is a diagonal matrix meaning that the residuals are uncorrelated. Given these assumptions and the above model, the variance-covariance matrix of  $x$  is written as

$$D(x) = L\Delta L' + \Psi.$$

Since  $L$  and  $F$  are both unknown, without losing generality, it can be assumed that  $D(f)=I$ , an identity matrix, and then

$$D(x) = LL' + \Psi. \quad (4.36)$$

This formulation together with the model in Equation (4.35) forms the standard factor model. In addition, since  $\Psi$  is a variance-covariance matrix,  $L$  must be a matrix satisfying the property that  $L'\Psi^{-1}L$  is a diagonal matrix with positive elements.

The factor model in Equation (4.35) is not unique in that the same covariance structure of  $D(x)$  may result from two different combinations of  $L$  and  $f$ . Any orthogonal transformation to the factor loading matrix  $L$  is equivalent to a rotation of the axes and the communalities or total variations are not affected by factor rotations. This may possess an advantage. Since all factor loadings are related to each other via some orthogonal transformation, a particular factor loading matrix can be chosen to possess some other meaningful properties. In fact, estimated factor loadings are usually rotated using an orthogonal transformation to help interpretation.

Solving a factor analysis problem appropriately is difficult. Many different approaches have been



suggested, including both non-iterative methods such as the Principal Component Method (Hotelling, 1933; Rao, 1964), the Principal Factor Method, Image Analysis (Guttman, 1953), and Canonical Factor Analysis (Rao, 1955), and iterative methods, such as the Maximum Likelihood Method (Lawley 1940), the Unweighted Least Squares Method (Joreskog, 1977), the Iterated Principal Factor Method, and Alpha Factor Analysis (Kaise and Caffrey, 1965). Simulation studies have shown that when the sample size and number of variables are large, all the methods appear to yield similar solutions (Browne, 1968; Linn, 1968; Velicer, 1977; Acito and Anderson, 1980).

After the factor loadings are determined by one of the above methods, if they are not readily interpretable, it is customary to transform them using orthogonal rotations so that a meaningful interpretation can be obtained with the new factor loadings. Usually, factors are rotated to make only a few variables have very large absolute loadings while the rest of the variables receive a small or zero loading. This makes the models easier to interpret as the factors can be interpreted as a linear combination of only a few variables. If a common characteristic exists in these few variables, the corresponding factor can then be interpreted as representing this common characteristic.

Several methods of orthogonal rotation have been proposed including the Quartimax Rotation (Neuhaus and Wrigley, 1954), Raw Varimax Rotation (Kaiser, 1958), Varimax Rotation, and Harman's Rotation method (Harman, 1976). However, the resulting factors may still not be interpretable after an orthogonal transformation is made. In this case, some non-orthogonal or oblique rotation methods have been developed, such as the HK rotation (Harris and Kaiser, 1964) and the PROMAX rotation (Hendrickson and White, 1964). These rotations lead to correlated factors so that the new transformed axes of the correlated factors pass through the clusters of the factor loadings more closely than those of orthogonal transformations.

The literature discussed in Chapter 2 suggests that there may exist four to five different aspects of personal risks to adverse birth outcomes. By carefully selecting and adjusting methods for calculating the factor loadings and rotation methods, these distinct aspects can be extracted to represent personal risks of adverse birth outcomes.

## **4.5 Summary**

This chapter has examined current global and local spatial statistical modeling techniques that are relevant to the proposed analyses of adverse birth outcomes. This discussion has laid a sound mathematical foundation for analyzing the social and spatial determinants of adverse birth outcomes. Based on the discussion of available methods, appropriate models were selected for the proposed analyses and hypothesis testing, which is discussed in the following chapter. While the majority of the discussion in this chapter has focused on reviewing relevant spatial and multilevel modeling

methods and spatial clustering analysis techniques, secondary analysis methods that were necessary for the proposed analysis were also reviewed, including geocoding, spatial interpolation and factor analysis. Based on the theoretical assumptions given in Chapter 3 and the methodologies provided in this chapter, an empirical study is conducted using data obtained from three health units in Ontario, Canada. The next chapter discusses the detailed analysis, results and interpretation from this study.

## **Chapter 5**

### **Case Study**

To test the proposed hypotheses and fulfill the thesis objectives, this chapter provides a detailed empirical analysis of the spatial and social impacts of adverse birth outcomes (LBW and preterm births) within a set of study areas in southern Ontario, Canada. The analysis is focused on available data in three public health unit areas, namely the Wellington-Dufferin-Guelph Public Health Unit (WDG), the Halton Region Health Unit and the Windsor-Essex County Health Unit (WEC). After personal and neighbourhood level indicators are derived from the available data, the statistical methods identified and explained in the previous chapter are used to conduct both global and local analyses to test the proposed five hypotheses presented in Chapter 3 and map the spatial variation of adverse birth outcomes and personal-level impacts. The results are then interpreted to explain and address the debates raised in the hypotheses. The areas that are most vulnerable to adverse birth outcomes and the spatial variation of personal risks are identified. The chapter concludes with a discussion of possible health policy improvements and community-based health interventions to reduce adverse birth outcomes and their inequalities.

#### **5.1 Description of the case study setting**

##### **5.1.1 Geographic and demographic context of the study region**

The spatial locations of the three health units under investigation and major cities within these health units are shown in Figure 5.1. The WDG and Halton regional health unit areas are located in the south central part of the province of Ontario, Canada, while the WEC health unit is located in the south west part of the province. The areas covered by the three health regions are situated astride Canada's most important and heavily-populated transportation corridor (Highway 401). Windsor, the southernmost city in Canada, lies at the western end of this corridor. It is situated across the Detroit River and Lake St. Clair from the city of Detroit, Michigan, in the United States.

Based on Canadian Census data, the population of Windsor in 2006 was 216,473. Visible minorities comprise 21% of the population, the fourth-highest percentage among Canadian cities, making it the most diverse city in Ontario outside of the Greater Toronto Area (GTA) (Statistics Canada, 2006). Although administrated separately, Windsor is within Essex County, Ontario. Essex County is one of the most agriculturally productive counties in Canada, and is the second most populated county in Ontario. The population of Windsor and Essex County, including the townships

of Amherstburg, Essex, Kingsville, Lakeshore, LaSalle, Tecumseh, and Leamington, was approximately 393,402 in 2006. The overall proportion of visible minorities in the region in 2006 was about 14%. Median income in 2005 for all census families in this region was \$71,605. Ten percent of the population in this region had low after-tax incomes in 2005.

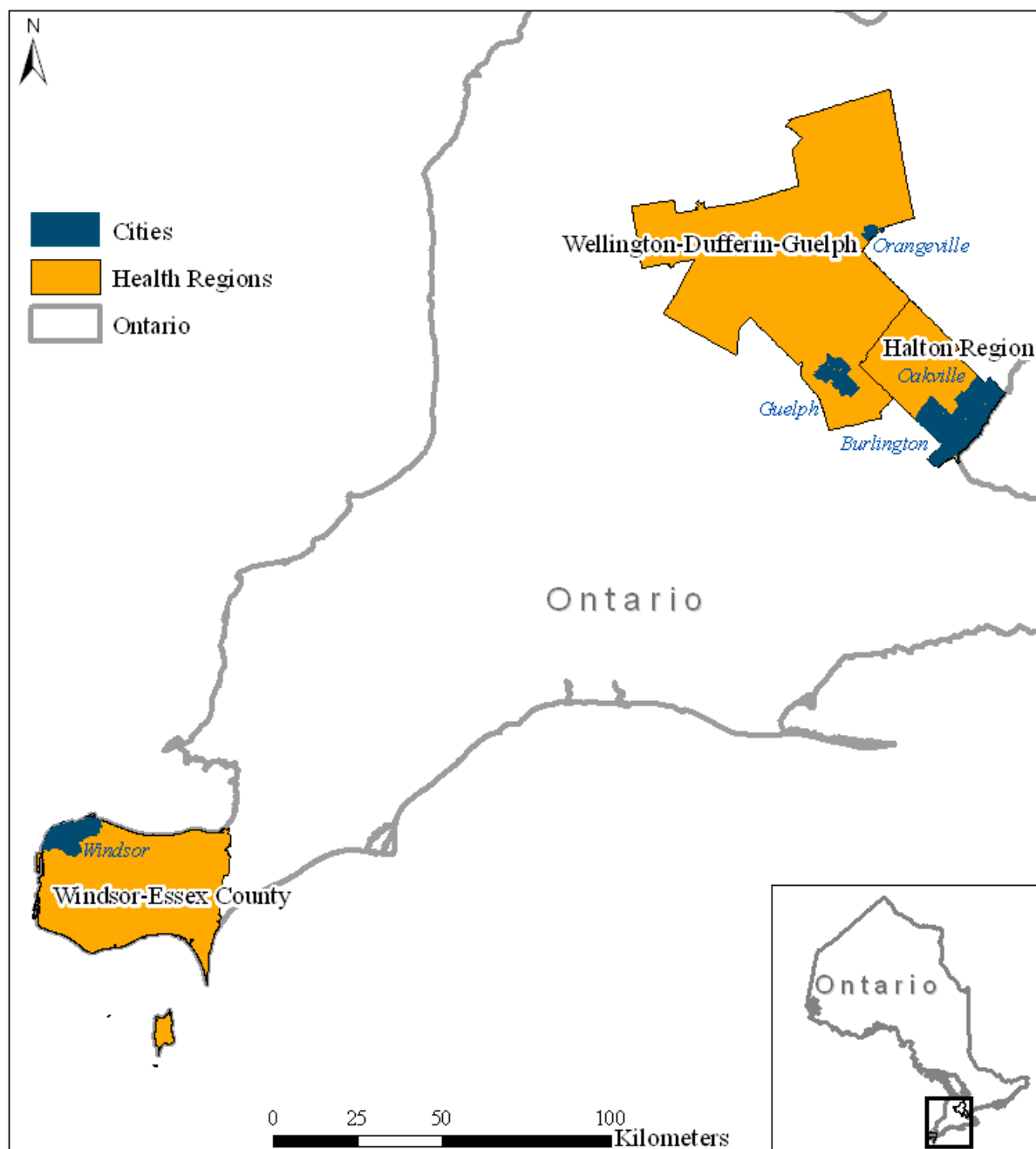


Figure 5.1: Spatial locations and major cities of the three health units

WDG and Halton regions are located about three hundred kilometres north-east of Windsor and Essex county. The WDG area incorporates 16 municipalities. These areas have a balance of

agriculture, commerce, industry, tourism and education. According to Canadian Census 2006 data, WDG had a population of 254,861. Some 7.6% of the population described themselves as visible minorities. Median census family income was \$75,667 in 2005 and about 6% of the population had low after-tax incomes.

Halton region is located in the southwest part of the Greater Toronto Area (GTA), which is the largest commercial, distribution, financial and economic centre in Canada and the third largest financial centre in North America. Halton comprises the city of Burlington and the towns of Oakville, Milton, and Halton Hills. It had a population of 439,256 in 2006, thirteen percent of which were visible minorities. Median census family income was 92,416 in 2005, the highest among the three health regions examined. Persons with low after-tax income comprised 6.5% of the population. Overall, the study region contains a diverse demographic, socio-economic, ethnic, and environmental composition, making it very appropriate for the analysis of adverse birth outcomes.

### **5.1.2 Current health practices of the study region**

The Canada Health Act (CHA), introduced in 1984, requires universal coverage for all insured persons for all medically necessary hospital and physician services, without co-payments or user fees. Currently, Canada operates a universal single-payer health care system, which collects medical fees from citizens and legal residents and pays for all services through a single government source. This system covered about 70% of total health care expenditures in 2005, which consumed 10% of the GDP in Canada. About 91% of hospital expenditures and 99% of total physician services were financed by the public sector at this time (Canada Institute for Health Information, 2005). Prescription drugs, dentistry and vision care continue to account for most of the private expenditures.

Canada's health care system is a kind of social insurance system, which provides public coverage for private delivery. Most services, including doctors, are delivered by private providers, although their revenues are primarily obtained from government billings. Canadian hospitals are controlled by private boards or regional health authorities, rather than being part of the Federal or Provincial Governments.

Although the CHA is Canadian Federal legislation, it deals only with how the system is financed. It does not address issues on how care should be organized and delivered. In Canada, health care delivery is within the jurisdiction of the provinces. Health regions, which are used to administer public health to Canadians, are organized by and given responsibilities from the provincial level of government. There are 36 health regions in the Province of Ontario. The study area examined for this thesis contains 3 of these 36 health regions. Within each health region, a public health unit (PHU) is established to administer health promotion and disease prevention programs. While each health unit is

governed by a Board of Health made up of elected representatives from the local municipal councils, the expenditures are shared by the Ontario Ministry of Health and Long Term Care and the constituent municipalities.

Traditionally, provincial governments across Canada fund health care services directly to delivery organizations, such as hospitals and community health centres. During the mid-1990s, this model of direct delivery was changed. Based on the view that local communities understand better their local health situation and associated risks, and are therefore better at determining their social service needs than centralized bureaucracies, provincial governments begin incrementally to devolve the responsibilities for allocating resources in health care service delivery to the community level. Based on the Local Health System Integration Act introduced in 2006, regional Health Authorities and local networks started to be established by provincial governments to be responsible for developing comprehensive service delivery programs.

In 2006, the Ontario Ministry of Health and Long-Term Care established a new division and divided the province into 14 regions of Local Health Integration Networks (LHIN). While not directly providing health services, each LHIN is responsible for planning, funding and integrating health care services for its own local communities, including hospitals, community care access centres, community support services, long-term care, mental health and addiction services, and community health centres. To develop integrated health service plans and service provision programs that tailor local needs and priorities and meet the provincial strategic directions, each LHIN consults with the stakeholders in its local communities, including the general public, patients, advocates, and health service providers, so that care can be better coordinated and more efficient.

Since health regions and LHINs operate independently of one another, their health service modes of delivery are different in approach and methodologies, which may cause differences in health outcomes between regions. Despite the differences, many similarities exist due to provincial-level regulation, programs and administrations. In terms of health promotion and prevention interventions for adverse birth outcomes, the prenatal care services provided by the three health regions under investigation are similar. They run the same provincial program, called “Healthy Babies Healthy Children” (HBHC), which is a prevention and early intervention initiative designed to help families promote healthy child development and help their children achieve their full potential. In terms of prenatal care, it uses screening and assessments to detect families who may be at risk and in need of support.

The screening tool used in this program consists of three questions to elicit responses on three factors that are considered predictors of at-risk pregnant women, namely the mother’s education level, prenatal class attendance, and smoking during pregnancy. The screening is conducted during

early pregnancy at prenatal clinics, physicians' offices or other community sites or by phone. After the scores of the above three questions are integrated and if a mother is identified as having potential risk, she will receive a brief assessment followed by an in-depth assessment to determine whether she or her family needs to receive prenatal components of the HBHC program. Depending on the needs of the at-risk mother and her family and the resources of the health unit, she will be given information, milk/food coupons, and home visits. She may also be referred to other prenatal care or specific services, such as the Canadian Prenatal Nutrition Program, Better Beginnings, Better Futures, substance abuse programs, smoking cessation programs, and other services in the community. These programs combined with regular prenatal care compose the prenatal care interventions in the three health regions. These programs are basically person-based, focusing mainly on medical care, prenatal education, mother's socioeconomic status, family conditions and social support, mother's nutrition, and pregnancy health behaviours.

An information system, called the Integrated Services for Children Information System (ISCIS), is used to record and administer the screening and assessment information that the HBHC program gathers about families, the services provided to families, and referrals to other services and programs. This ensures that the personal-level data collected across regions are consistent. Although collected and administered separately by corresponding health units, the live birth data in the ISCIS within the three health units examined have the same content and format, and therefore are able to be integrated. This provides a great opportunity to obtain desirable sample sizes to produce generalizable analysis findings.

## **5.2 Data Sources**

Data availability is a major obstacle to micro-level secondary health research. To obtain necessary personal-level live birth data, almost all the PHUs in central and south-west Ontario were contacted. Three of the health units, namely WDG, Halton, and WEC, showed great interest in the proposed research. The collaboration among the three health units, the University of Waterloo, and Cancer Care Ontario (CCO) allowed the adverse birth outcome research to take place and to help promote small-area analysis across health units using the Rapid Inquiry Facility (RIF) tool. The RIF tool provides a convenient approach to disease mapping and gives rapid impressions of the spatial distribution of disease incidences. Although it is not used in this thesis, the tool is useful for environmental health surveillance to identify potentially elevated rates of disease around environmental sources or in specific areas.

Personal live birth data were collected. Although the disjoint geographical location of WEC to the other health regions may pose some challenges for spatial analysis, the obtained overall sample

sizes (about 90,500 geocoded live births) are of considerable size and make the analysis results more powerful for revealing the relationships between adverse birth outcomes and various social and environmental risks in socially advanced societies.

At the contextual level, community- or neighbourhood-level data were obtained mainly from the 2001 and 2006 Canadian Census of Population and Dwellings and the Canadian Community Health Surveys (CCHS). The following sections provide further discussion on data acquisition and neighbourhood-level variable derivations.

### **5.2.1 Personal-level data extraction and management procedures**

The personal-level live birth data of the three health units were accessed through the ISCIS system. Both adverse birth outcomes and associated personal-level risk factors were obtained from the Postpartum Screening (PARKYN) sub-database. Postpartum screening uses a PARKYN tool that consists of a series of questions designed to identify factors associated with risks of parenting problems. The screening is administered before mothers leave hospital. Data are routinely collected and maintained in the PARKYN sub-database. The purpose of the postpartum screening is to identify mothers and babies who may be at risk and to link them to the available services. Although the aim of the screening is to reach all women who give birth in Ontario, data can only be collected for mothers who give birth at hospitals, which constitute approximately 98 percent of total live births.

The variables collected include birth weights, gestational ages, birth order, baby's gender, mother's demographics and health conditions (age, language, abortion, health challenge, infection, family history of genetic health challenge, evidence of schizophrenia, mental challenge, depression, and stresses related to delivery), mother's health behaviours (smoking, alcohol drinking and other drug uses, and maternal under-nutrition), and family's social situations and medical challenges (social support, lone parents, financial difficulties, no prenatal care, low education, failure to thrive, family violence, father's evidence of schizophrenia, father's mental challenge, and not attending prenatal classes). In addition, the residential addresses of mothers and the associated postal codes were also obtained to geocode the locations of all live births in the study areas. Although the extracted personal-level variables are not exclusive, they include most of the identified personal-level risks discussed in Chapter 2.

Since the ISCIS database contains personal information, procedures were taken to extract only the information necessary to conduct the proposed research. The data extraction and data management processes were controlled by each PHU to ensure data security. The major procedures used are summarized as follows:

- All of the required personal-level birth data mentioned above were extracted within the PHU



offices using the tools provided by ISCIS. Except mother's address, all the personal identifiers, such as names and contact information, were stripped from the extraction. Although individual addresses are considered to be personal identifiers, they were extracted from the database because they are necessary for locating the spatial locations of live births so that spatial multilevel analysis could be conducted to assess the spatial impact on adverse birth outcomes.

- For security reasons, the extracted data were partitioned into two data sets, namely a de-identified data set that contained only variables of adverse birth outcomes and personal risk factors, and an identity-only data set that contained the addresses. These two data sets had the same unique identifiers so that they could be linked whenever necessary.
- The identity-only data set was kept by the PHU. The geocoding process that used the addresses to generate spatial points of live births was undertaken within the health unit. After the geocoding was done, only the spatial point data with associated necessary personal risk factors of adverse birth outcomes were taken away from the health unit for analysis purposes.
- The removed data set was maintained and analyzed in the Statistics Canada's Research Data Centre at the University of Waterloo, which has a high standard level of data security in place.
- The release of the findings (including maps, graphs, tables, and writing documents) does not include any information at the personal level. All findings are presented only at aggregated levels (census dissemination areas and higher).

#### 5.2.1.i Geocoding

Using the geocoding by addresses method described in Chapter 4, the individual live births were first geocoded using home addresses. The "US Streets with zone" address locator style was used to find the location of each live birth on a specific side of the street. A newly updated Ontario Single Line Street Network (SLSN) in Environmental Systems Research Institute (ESRI) shapefile format, that includes all necessary variables (street numbers, street names, types, prefix direction, prefix type, and suffix type), was used as the reference. Due to missing addresses from the live birth database and the inconsistency between the SLSN and the home addresses, approximately 80-85% of all live birth records were geocoded correctly. This was deemed to be a successful result.

The remaining live birth records were then geocoded using the alternative method of geocoding by postal code. The PCCF+ files were used to match the 6-digit post code of each live birth to a pair of spatial coordinates and associate live birth points with corresponding census dissemination areas (CDA). Samples with the same postal code were randomly assigned to one of their associated dissemination areas based on the proportion of the population of these areas to avoid false peak values in one dissemination area.

Through the above two geocoding processes, over 98.5% of all live births records in the database were effectively geocoded, providing an excellent basis for the subsequent analysis. The rest of the records had either inaccurate street addresses or non-available postal codes. The geocoded spatial points were then used to conduct multilevel and spatial analysis to identify the social and environmental determinants of adverse birth outcomes.

### **5.2.2 Contextual-level data collection**

Contextual determinants operate at different spatial scales (neighbourhood, municipality, and higher scales). For the purpose of this thesis, neighbourhoods and municipalities were selected as the units of analysis. Due to a lack of definition and division, neighbourhoods are often represented by political boundaries, such as census dissemination blocks, CDAs, census tracts, and postal code areas, or planning divisions. The selection of an appropriate scale to represent neighbourhoods may be determined by the homogeneity of socio-economic conditions of spatial units at this scale, perceived sense of neighbourhood by the local residents, and the zones of influence of community organizations. It may also be statistically determined by the distance at which the greatest spatial clustering of adverse birth outcome instances and the clustering of socio-economic status are obtained. At a higher spatial scale, the rural-urban division and the division of different municipalities are “natural” divisions, since health-related policies may operate differently from municipality to municipality. These higher scale units may be appropriate administrative units for the analysis of policy impacts on adverse birth outcomes as well. At the smaller spatial scale, since CDAs are the most homogenous spatial divisions with appropriate sample sizes and spatial coverage, they were used as a surrogate of neighbourhoods for the analysis. Neighbourhood-level data were then collected accordingly.

Based on the period of the live birth data collected, the socio-economic and demographic characteristics of neighbourhoods were obtained from the corresponding Canadian Census of Population and Dwellings (2001 and 2006) at the CDA level. Variables that have potential impacts on adverse birth outcomes were extracted, including various socio-economic status variables (incomes, education, and employment), ethnicity (visible minority, non-Canadian citizen, immigrant, aboriginal, and non-official-language-speaking), and living conditions (dwelling values, major house repair, overcrowding, rent rate, and residential instability).

The other data source used to obtain community-level data was the Canadian Community Health Surveys (CCHS). The CCHS is a national-wide cross-sectional survey that collects information related to health status, health care utilization and health determinants for the Canadian population. Prior to the year 2007, data were collected every two years. The first three cycles for the years of

2001, 2003 and 2005 were available and used for the thesis analysis. Since 2007, data are collected every year with the goal of improving its effectiveness and flexibility. The surveys include common content ranging from Alcohol, General Health, Health Care Utilization, Exposure to Second-hand Smoke, through Physical Activities and Income; and optional content, such as Access to Health Care Services, Stressors, Depression, Food Security, Health Care System Satisfaction, Home Safety, Satisfaction with Life, Self-Esteem, Social Support, Voluntary Organizations, and Work Stress. All variables may have potential uses for the analysis of adverse birth outcomes in this research. However, since the data are collected for community health research at the health region level, they are not suitable for the required small-area analysis at the neighbourhood level. Hence, spatial interpolation procedures were used to derive neighbourhood variables.

#### 5.2.2.i Spatial interpolation

Based on the literature reviewed in Chapter 2, variables in the CCHS data with potential influences on adverse birth outcomes were identified and extracted, including Self-Perceived Health (SPH), Chronic Health Conditions (CHC), Self-Perceived Unmet Healthcare need (SPUH), Self-Perceived Stress (SPS), Sense of Belonging to local Communities (SBC), Emotional Unhappiness (EU), Food Insecurity (FI), Insufficient Vegetable Intake (IV), Daily Smoking (DS), Smoking Inside Home (SIH), Physical Inactivity (PI), Regular Drinking (RD), and Heavy Drinking (HD). In addition, Household Income (HINC) was also extracted from the CCHS data so that it could be compared with the readily available CDA-level Census data (2006) to assess the accuracy of the interpolation results.

As described in Chapter 4, spatial interpolation methods are based on the assumption of spatial dependency. It is necessary to assess spatial autocorrelation before spatial interpolation is used. The Global Moran's I method defined in Equation (4.18) was used to assess the spatial dependency among CCHS variables. Given a set of sample locations and an associated attribute, Moran's I evaluates whether the pattern expressed is clustered, dispersed, or random. In general, a Moran's I index value near 1.0 indicates perfect clustering while an index value near -1.0 indicates perfect dispersion. The provided p-value also indicates whether or not the calculated index values are statistically significant or, in other words, whether or not the observed autocorrelations indicate truly non-zero dependence, or whether they are have just occurred by chance.

The CCHS samples obtained from the Windsor and Essex County (WEC) health region were used to examine the spatial dependency and demonstrate the effectiveness of different interpolation methods. An Inverse Distance Squared weighting scheme was applied to calculate the Moran's I values for selected CCHS variables. The neighbourhood search threshold was 4652m, which is the largest distance between any two neighbouring units in WEC. Table 5.1 lists the results.

<i>CCHS variable</i>	<i>Moran's I Index</i>	<i>Variance</i>	<i>Z Score</i>	<i>p-value</i>
<b>Household income (HINC)</b>	<b>0.1189</b>	<b>0.0002</b>	<b>9.4170</b>	<b>0.0000</b>
<b>Self-perceived health (SPH)</b>	<b>0.1020</b>	<b>0.0002</b>	<b>8.0648</b>	<b>0.0000</b>
<b>Chronic health conditions (CHC)</b>	<b>0.0302</b>	<b>0.0002</b>	<b>2.4028</b>	<b>0.0163</b>
Self-perceived unmet health need (SPUH)	0.0053	0.0002	0.4250	0.6708
Self-perceived stress (SPS)	-0.0080	0.0001	-0.8627	0.3883
<b>Sense of not belonging to local communities (SBC)</b>	<b>0.0248</b>	<b>0.0002</b>	<b>1.9737</b>	<b>0.0484</b>
<b>Emotional unhappiness (EU)</b>	<b>0.0633</b>	<b>0.0001</b>	<b>7.1493</b>	<b>0.0000</b>
<b>Food insecurity (FI)</b>	<b>0.0985</b>	<b>0.0002</b>	<b>7.7958</b>	<b>0.0000</b>
<b>Insufficient vegetable intakes (IV)</b>	<b>0.0308</b>	<b>0.0002</b>	<b>2.4449</b>	<b>0.0145</b>
<b>Daily smoking (DS)</b>	<b>0.0266</b>	<b>0.0002</b>	<b>2.1218</b>	<b>0.0339</b>
Smoking inside home (SIH)	-0.0082	0.0002	-0.6255	0.5316
<b>Physical inactiveness (PI)</b>	<b>0.0300</b>	<b>0.0002</b>	<b>2.3831</b>	<b>0.0172</b>
<b>Regular drinking (RD)</b>	<b>0.0334</b>	<b>0.0002</b>	<b>2.6572</b>	<b>0.0079</b>
<b>Hard drinking (HD)</b>	<b>0.0312</b>	<b>0.0002</b>	<b>2.3364</b>	<b>0.0195</b>

Table 5.1: Moran's I results for selected CCHS variables (Bold numbers represent statistically significant results at the 95% confidence level)

The p-values in Table 5.1 indicate that most of the CCHS variables in WEC are spatially auto-correlated. Although the autocorrelations are moderate, they do show statistical significance. The existence of these spatial dependencies for most of the CCHS variables supports the use of spatial interpolation methods for estimating values at unknown locations. Several variables, including SPS, SIH, and HD do not show significant spatial dependencies globally. Hence, the use of spatial interpolation methods on these variables is questionable.

As mentioned in Chapter 4, the confirmed CCHS spatial dependence may be due to a spatial process that applies universally across the study region, for example, distance impacting on the interaction and communication of local residents. In this situation, the Kriging method may be a better choice for spatial interpolation. On the other hand, the spatial clustering may be a result of local constraints, such as the clustering of housing conditions caused by local zoning regulations. In this case, IDW may be a better choice. Both Kriging and IDW were investigated using the spatial interpolation tools in the Geostatistical Analyst extension in ESRI's ArcGIS software. The household income variable in CCHS data and the same variable obtained from the Census 2006 were compared to examine the effectiveness of these methods.

Table 5.2 lists selected cross-validation results with different methods and parameters for interpolating the household income variable. It can be observed that Kriging generally performed better than IDW methods in terms of the root-mean-square prediction errors. This suggests the existence of underlying universally applied rules of spatial dependence. By carefully adjusting the Kriging parameters, the fitted semi-variogram function for the household income variable is produced in Figure 5.2.

<i>Spatial interpolation method and parameters</i>	<i>Mean prediction error</i>	<i>Root-Mean-Square</i>	<i>Average Standard Error</i>	<i>Mean Standardized</i>	<i>Root-Mean-Square Standardized</i>
Ordinary Kriging Exponential function lag: 200, number lags:100	-230.9	45360	41860	-0.0053	1.083
Ordinary Kriging Exponential function lag: 100, number lags:100	-280.5	45360	43190	-0.0063	1.05
Ordinary Kriging Exponential function lag: 200, number lags:50	-279.2	45360	43160	-0.0063	1.051
Ordinary Kriging Spherical function lag: 200, number lags:100	-295.9	45400	42580	-0.0068	1.066
IDW with power 2	-906.4	50680			
IDW with power 1	-1060	48110			

Table 5.2: Cross-validation comparison between models and parameters

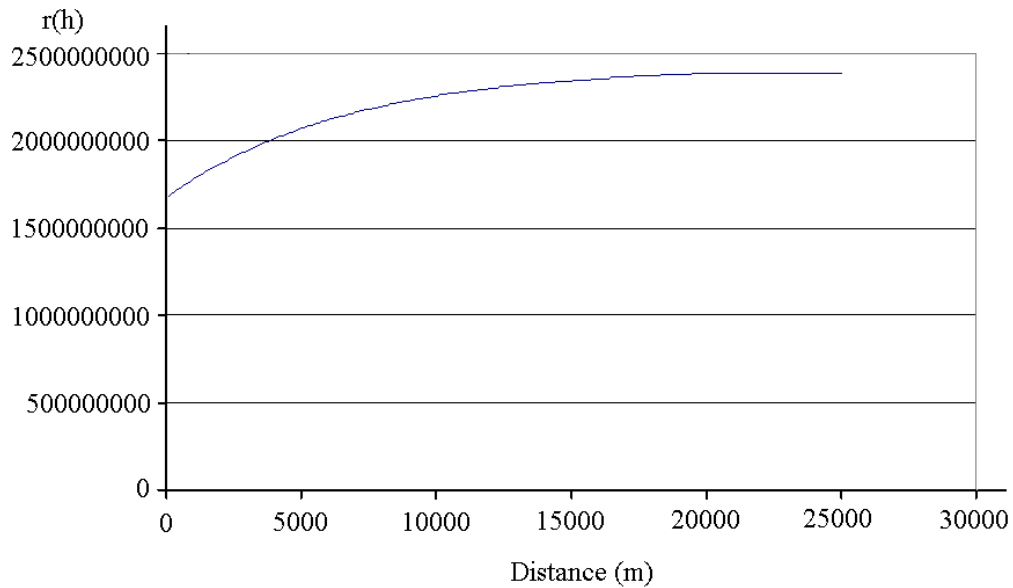


Figure 5.2: The best fitted semi-variogram function for household income

The semi-variogram shows that the spatial dependency is largest at a small distance. The variogram values gradually increase with the increase of distance and reach the sill at about 20 kilometres. To ensure that short-range autocorrelation is not masked, a small lag size (200m) was selected in this situation so that the fitted function may better represent the real distance-dependent spatial autocorrelation. Subsequently, 100 lags were selected to cover the range of the semi-variogram. Due to small ranges for most of the CCHS variables, spatial dependencies only exist at small spatial distances. It was therefore deemed to be better to use small lag sizes and a large number of lags to cover the ranges of the semi-variogram functions. Depending on the spatial structures of different CCHS variables, lag sizes between 50m and 200m were selected for the interpolation of variables of interest. The exponential and spherical functions showed similarly good performance on the fit of the empirical semi-variogram. The selection of one over another was determined by comparing their corresponding root-mean-square prediction errors. In the case of household income interpolation, as shown in Table 5.2, the exponential function fitted a little better than the spherical function and produced a relatively smaller root-mean-square error.

The selection of neighbourhoods is also affected by the small ranges of the semi-variogram. A small range in the semi-variogram means that spatial dependencies only exist at small distances. A large number of neighbourhoods may not help to improve interpolation accuracy, but may produce an overly smoothed surface, which is not expected for values at the small-area level. A suitable approach to address this is to use samples from the same neighbourhood or surrounding neighbourhoods for interpolation so that local characteristics can be better preserved. Since no directional effect was found by comparing different neighbourhood structures using ellipse and circle shaped neighbourhoods and dividing neighbourhoods into sectors, a fixed set of 50 nearest neighbours based only on Euclidean distances was finally chosen to make sure enough sample points were used for interpolation at each unknown location, and no unnecessary long-distance sample points were involved. Thus, local variations could be distinguished with a reasonable number of sample points.

The interpolated surfaces using IDW interpolation with power 1 and using Kriging with ideal parameters are plotted in Figures 5.3 and 5.4. To generate further CDA-level variables, these two surfaces were then aggregated by averaging the cell values in each CDA respectively. The results are plotted in Figures 5.5 and 5.6. It can be observed by comparing Figure 5.3 with Figure 5.4 (or comparing Figure 5.5 with Figure 5.6) that Kriging interpolation resulted in a relatively smoother surface than the IDW method due to its use of the universally applied semi-variogram for weighting. The IDW interpolation, on the other hand, is subject to local spatial clustering. Hot spots of high interpolated values can be observed from place to place in Figure 5.3 due to high values or outliers in these locations.

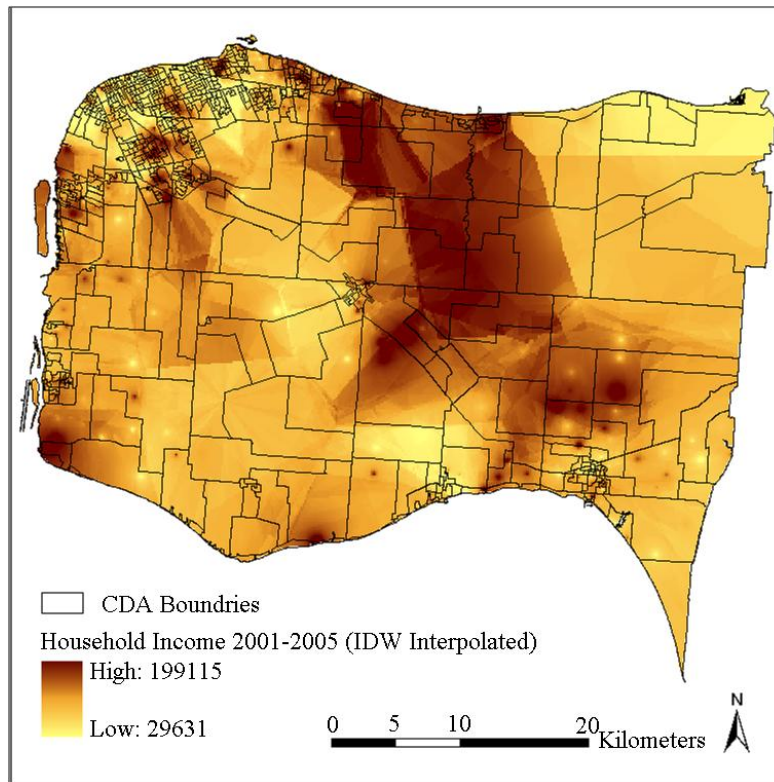


Figure 5.3: Household income distribution interpolated using IDW with power 1

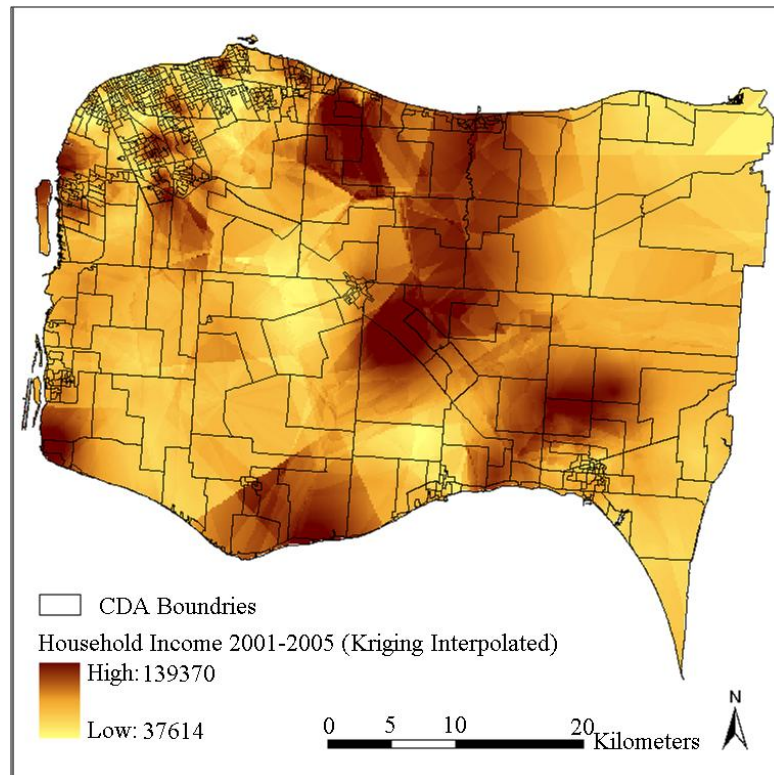


Figure 5.4: Household income distribution interpolated using Kriging

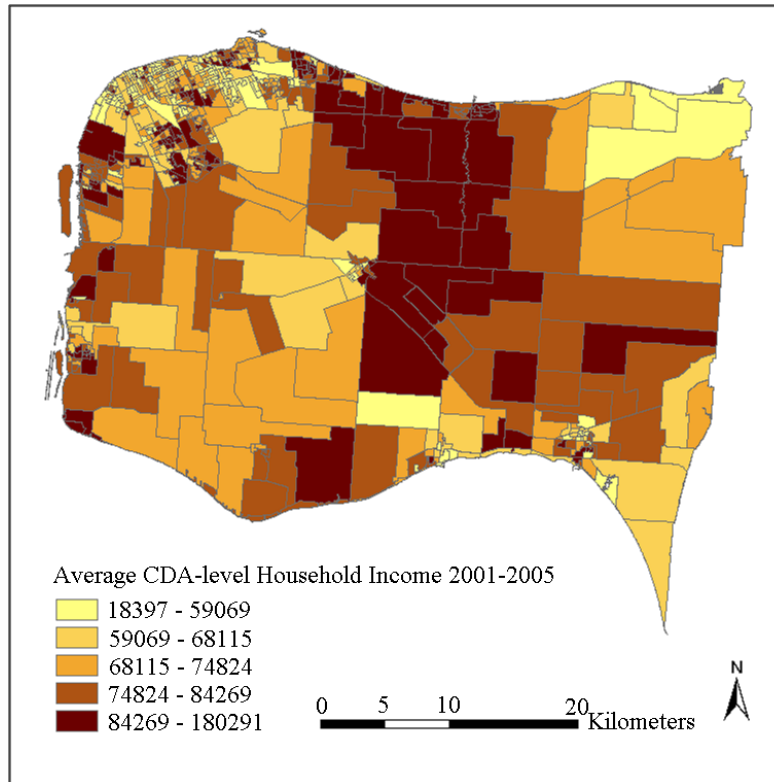


Figure 5.5: Aggregated CDA-level household income by IDW interpolation

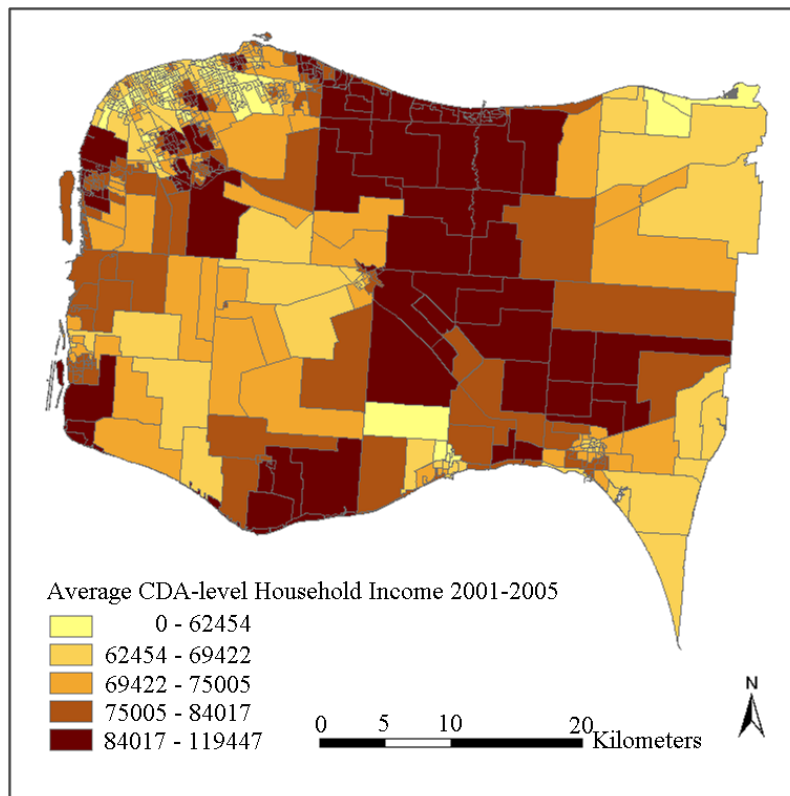


Figure 5.6: Aggregated CDA-level household income by Kriging interpolation



To demonstrate the effectiveness of the interpolated results on the representation of the real CDA-level status, CDA-level household income values were extracted from the 2006 Census (Figure 5.7) for comparison. The standard errors produced by the Kriging interpolation are also presented in Figure 5.8 to show the correctness and reliability of the interpolated results.

Although the CCHS samples were collected between 2001 and 2005 and the status of some CDAs may have changed after that, the overall neighbourhood status of the study region has remained relatively constant. Figure 5.7 shows quite similar overall spatial patterns as Figure 5.5 and Figure 5.6, in that household incomes are lowest in cities and towns and higher in some sectors of suburbs and rural areas. Differences can also be observed in that the census data show a more random pattern and some of the values are not matched with the interpolated results, especially in some rural areas with less population. For example, some CDAs at the upper-right and lower-right corners show inconsistent values. This inconsistency can be partially explained by the larger standard errors due to the lack of samples at these locations or edge effects of Kriging near the boundaries (Figure 5.8). This may also be due to some actual changes at the local area level, or data bias between the surveys and the census.

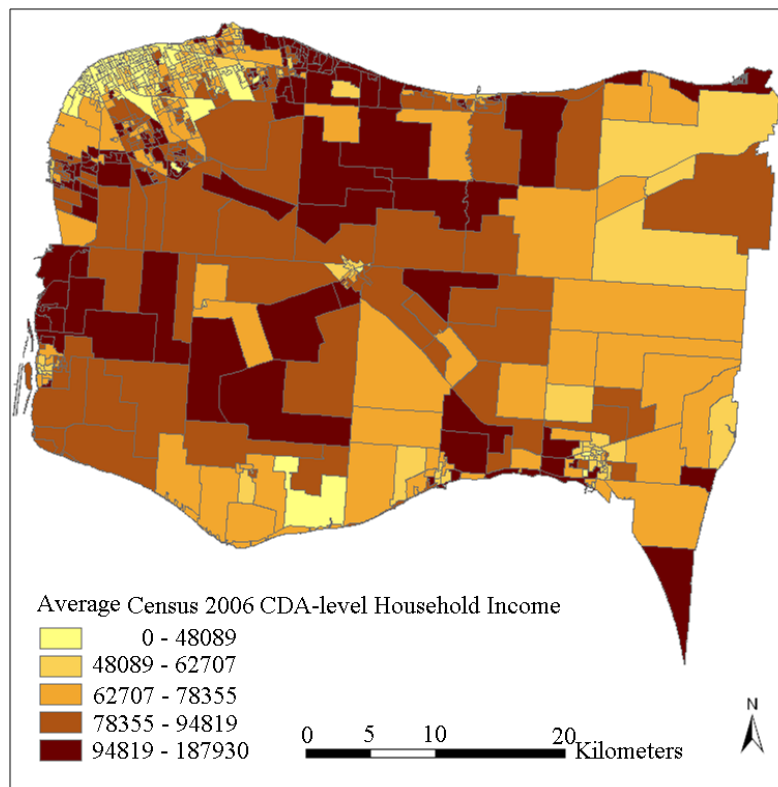


Figure 5.7: Census 2006 CDA-level household income

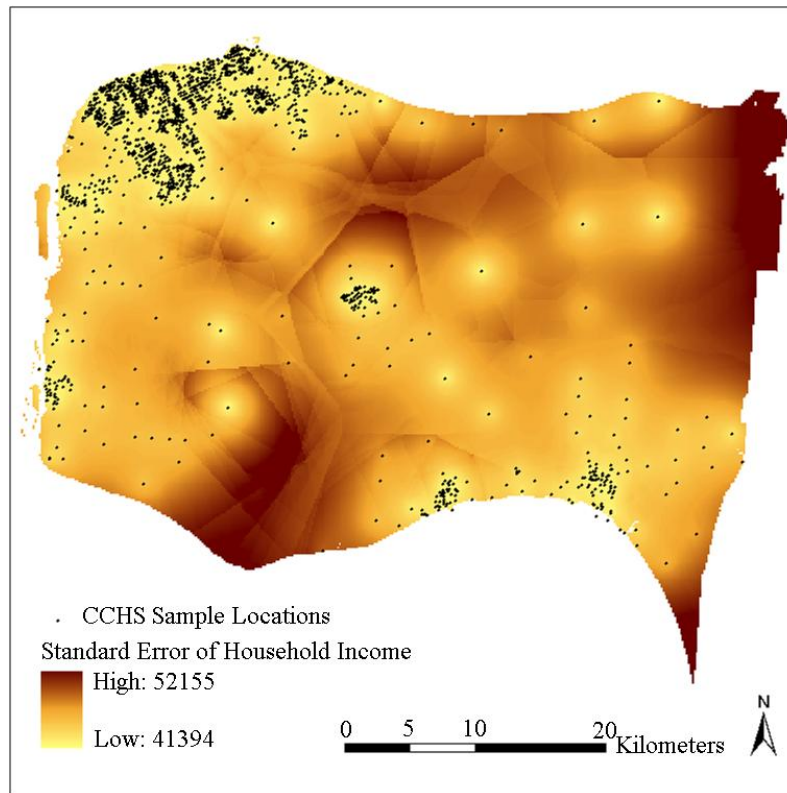


Figure 5.8: Kriging standard errors for the household income variable

Figure 5.7 shows that the actual spatial distribution of household income is determined by both global forces and local variance. The mean-squared-differences were also calculated between the Kriging result and the Census data, and between IDW results and the census data. Kriging still shows a slightly better performance than IDW. This is consistent with the results presented in Table 2. Nevertheless, the overall gradients and patterns, and the relative status remain consistent between the interpolated results and the Census 2006 data.

Using similar Kriging procedures, the variables in Table 5.1 were also interpolated and aggregated to the CDA-level for all the three regions to obtain neighbourhood-level indicators. These indicators were then used as explanatory variables in statistical modeling to assess neighbourhood determinants of adverse birth outcomes.

### 5.3 Adverse birth outcome Analysis

#### 5.3.1 Current adverse birth outcomes of the study region

Depending on availability in the ISCIS database, data gathered in the Wellington-Dufferin-Guelph Health Unit were from the years 2003 to 2008, and data in the Halton Region Health Unit and the Windsor-Essex County Health Unit were from the years 2000 to 2008. There were 91,796 live

births in total in the three health units in the study period, and 90,500 of them were successfully geocoded. Based on these live birth records, LBW and preterm birth rates of the three health units by years are reported in Table 5.3 and Table 5.4. The preterm birth rates for Ontario and Canada were obtained from the Public Health Agency of Canada (2008). The LBW rates for Ontario and Canada were obtained from Statistics Canada. They are also listed in these two tables for comparison.

<i>Health Units</i> <i>Year</i>	<i>Halton</i>	<i>Wellington- Dufferin-Guelph</i>	<i>Windsor- Essex County</i>	<i>Total</i>	<i>Ontario</i>	<i>Canada (excluding Ontario)</i>
<b>2000</b>	7.37	-	9.39	8.16	7.3	7.6
<b>2001</b>	7.2	-	5.97	6.60	7.1	7.5
<b>2002</b>	6.95	-	6.93	6.94	7.3	7.6
<b>2003</b>	6.28	6.44	7.77	6.88	7.4	7.9
<b>2004</b>	7.41	6.13	8.78	7.67	7.6	8.2
<b>2005</b>	7.02	6.52	8.46	7.50	-	-
<b>2006</b>	6.51	6.57	8.08	7.11	-	-
<b>2007</b>	7.33	6.58	8.58	7.65	-	-
<b>2008</b>	7.3	6.71	9.75	8.19	-	-
<b>Total</b>	7.03	6.48	8.17	7.41		

Table 5.3: Preterm birth rate by birth year and health units (%)

<i>Health Units</i> <i>Year</i>	<i>Halton</i>	<i>Wellington- Dufferin-Guelph</i>	<i>Windsor- Essex County</i>	<i>Total</i>	<i>Ontario</i>	<i>Canada</i>
<b>2000</b>	5.32	-	6.08	5.74	5.7	5.6
<b>2001</b>	5.29	-	5.48	5.39	5.6	5.5
<b>2002</b>	5.45	-	5.44	5.45	5.9	5.7
<b>2003</b>	5.44	4.76	5.45	5.29	6.1	5.9
<b>2004</b>	4.80	4.23	6.30	5.27	5.9	5.9
<b>2005</b>	5.12	5.19	6.07	5.51	6.2	6.0
<b>2006</b>	5.17	4.77	6.08	5.42	6.2	6.1
<b>2007</b>	5.42	4.86	6.32	5.63	6.2	6.0
<b>2008</b>	5.48	5.31	6.53	5.87	-	-
<b>Total</b>	5.27	4.82	5.98	5.50	5.98	5.84

Table 5.4: Low Birth Weight rate by birth year and health units (%)

The overall preterm birth rate is 7.41 and LBW rate is 5.5 for the three health units. The numbers are a little smaller than the overall rates for Canada in general. The Windsor-Essex County Health

Unit had the highest preterm births rate (8.17) and LBW rate (5.98). Since the year 2001, both preterm birth rates and LBW rates gradually increased. The rate of preterm births increased from 6.6 in 2001 to 8.19 in 2008 and the rate of LBW has been increased from 5.39 in 2001 to 5.87 in 2008. As discussed in Chapter 1, the increases of LBW and preterm birth rates may largely be attributed to greater use of assisted reproduction techniques and increased multiple births. Increasing social inequalities over this time period may also be the reason causing this increase. Further analysis is needed to test whether or not this increase is statistically significant.

The spatial distribution of the raw rates at the 2001 CDA level of preterm births and LBW are shown in Figures 5.9 and 5.10. While Figures 5.9 and 5.10 show some interesting locations of high preterm birth and LBW rates, it is not so easy to infer any overall patterns from the maps. To assess the spatial and social impacts on adverse birth outcomes, the five hypotheses stated in Chapter 3 are tested in the following sections.

### 5.3.2 Hypothesis 1: Personal vs. Spatial

The first step toward spatial or neighbourhood analysis is to assess whether or not spatial impacts exist and of so to what extent they are evident. As discussed in Chapter 4.2.3, the intra-class correlation coefficient (ICC) provides a direct quantitative measure of similarity between individuals within areal units. Among many different estimators for ICC, three are considered to be most accurate in terms of bias and mean square error (Zou & Donner, 2004), namely the FC estimator in Equation (4.11), the Pearson pairwise estimator in Equation (4.12), and the analysis of variance (ANOVA) estimator in Equation (4.13). They are used to quantify the extent of neighbourhood variations in LBW and preterm births in the study region. The results are shown in Table 5.5.

<i>Estimator</i>	<i>ICC results for singular preterm birth</i>	<i>ICC result for singular full-term LBW</i>
FC estimator	0.0037	0.0029
Pearson pairwise estimator	0.0014	0.0012
ANOVA estimator	0.0037 [0.0011,0.0062]	0.0029 [0.0005, 0.0051]

\*Numbers in square brackets are confidence intervals at 95% confidence level.

Table 5.5: ICC estimates for LBW and preterm birth

The results of the above three estimators show consistently small ICC values for both LBW and preterm births. Despite the small ICC values, the confidence intervals calculated by the ANOVA estimator indicate that the intraclass correlations are statistically significant. As discussed earlier in Chapter 4, when the intraclass correlations are small, these estimators tend to underestimate the ICC. Hence, in this case, it is difficult to interpret ICC as the proportion of total variance.

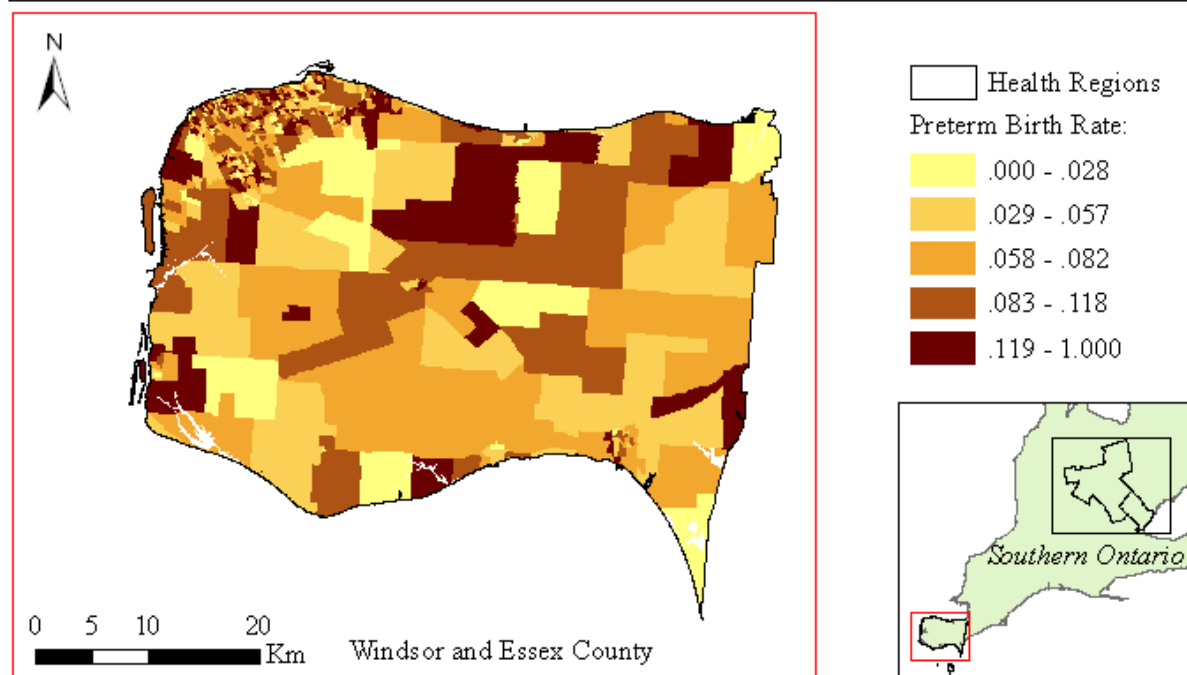
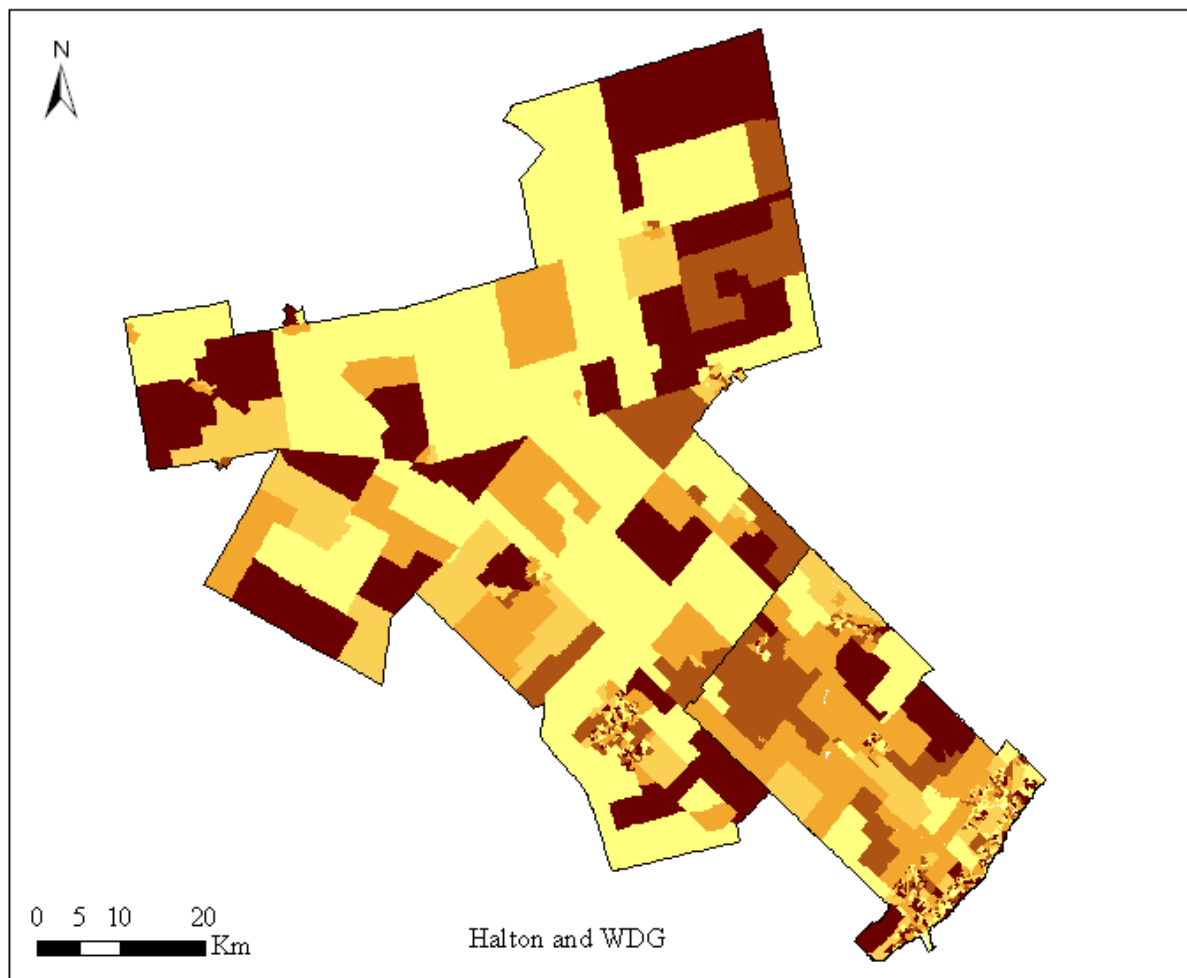


Figure 5.9: The spatial distribution of raw preterm birth rates

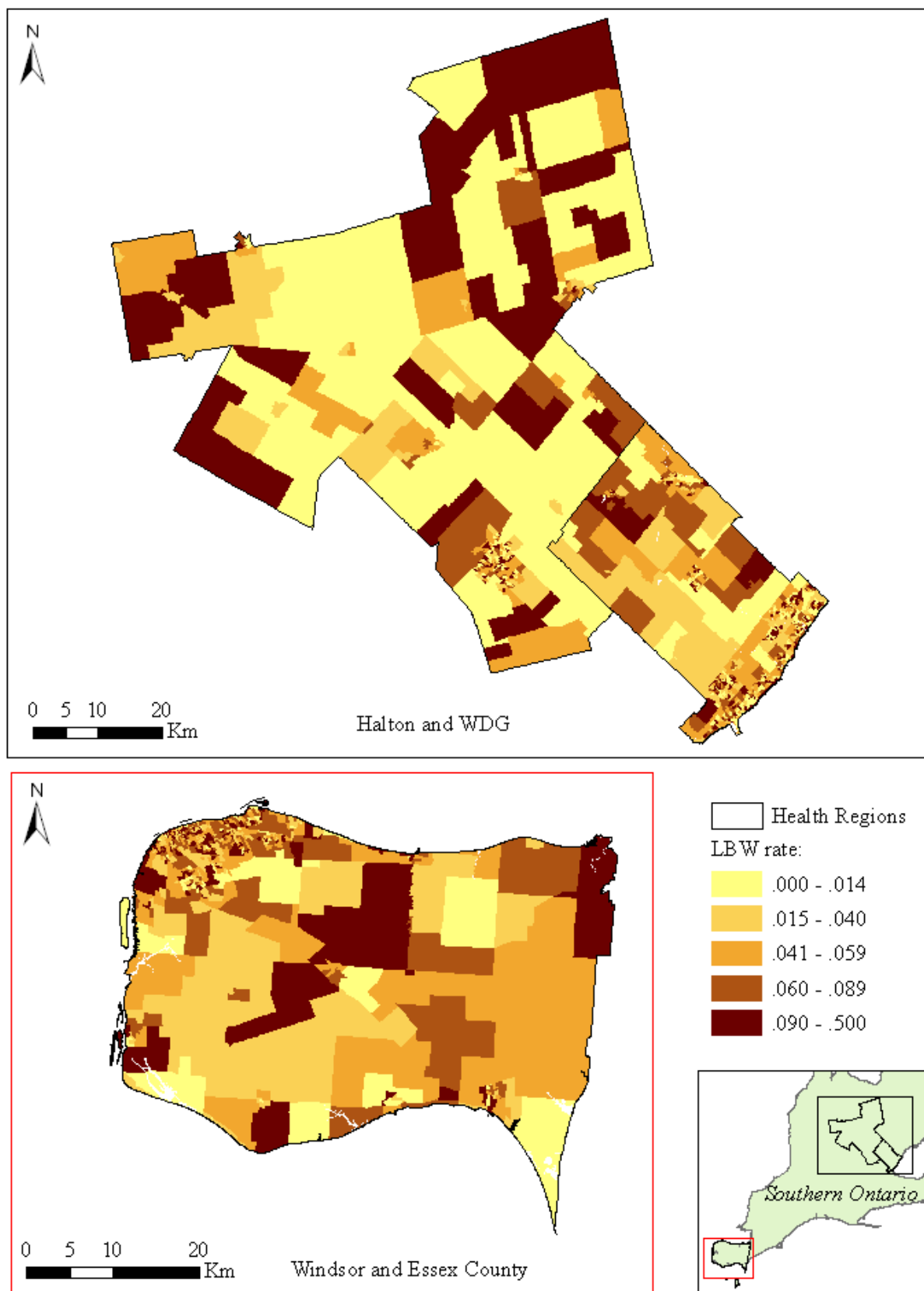


Figure 5.10: The spatial distribution of raw LBW rates

The generalized linear mixed model approach was then used to test the magnitude of neighbourhood variation. Since the adverse birth outcomes of interest are IUGR and singular preterm births, a *logit* model was constructed similar to Equation (4.14):

*Adverse birth outcomes* ~ *binary* ( $p_{ij}$ )

$$\text{Level 1 (personal): } \logit(p_{ij}) = \beta_{0j} + \beta_1 \text{AGE19}_{ij} + \beta_2 \text{AGE36}_{ij} + \beta_3 \text{FEMALE}_{ij} + \beta_4 \text{MLTIBIRTH}_{ij} \\ + (\beta_5 \text{PRETERMBIRTH}_{ij})$$

$$\text{Level 2 (neighbourhood): } \beta_{0j} = \gamma_{00} + v_{0j}. \quad (5.1)$$

This model controls mother's age, babies' sex, multiple births, (and preterm births when LBW is the response variable) at the personal level, so that the neighbourhood variations of standardized singular preterm births and singular full-term LBW births can be modeled. At the neighbourhood-level, only a fixed effect and a random term were included. A *probit* model can be similarly constructed simply by replacing the link function  $\logit(p_{ij})$  to  $\text{probit}(p_{ij})$ . After the neighbourhood variations  $v_{0j}$  were obtained, the ICC values were then be calculated using Equations (4.15) and (4.17) respectively. The *logit* model was estimated using both a doubly-iterative pseudo-likelihood technique (SAS GLIMMIX Procedure) (Wolfinger and O'Connell, 1993; Breslow and Clayton, 1993) and a Bayesian approach (WinBUGS), while the *probit* model was estimated by the Bayesian approach only. The estimated results are shown in Table 5.6.

The results in Table 5.6 show that the *logit* model produced slightly higher ICC values than the *probit* model. Overall, about 2-3% of the total variation in the dependent variables is attributed to neighbourhood variations. While these values are trustworthy in their calculation, the personal-level variances were calculated based on binary results (1 and 0 to represent whether or not there is a LBW

<i>Estimator</i>	<i>preterm birth</i>		<i>LBW</i>	
	<i>Variance of <math>v_{0j}</math></i>	<i>ICC</i>	<i>Variance of <math>v_{0j}</math></i>	<i>ICC</i>
Logit model (SAS GLIMMIX)	0.0739	0.022	0.1099	0.032
Logit model (WinBUGS)	0.0743	0.022	0.1139	0.0335
Probit model (WinBUGS)	0.0138	0.0136	0.0192	0.0188

Table 5.6: ICC estimates for LBW and preterm birth using random effect models

or preterm birth). This actually exaggerates the personal-level variance since the continuous birth weights (or gestational ages) are truncated into the two extremes. The neighbourhood-level variance may be underestimated due to the exaggeration of the personal-level variance.

The concept of design effect (Kish, 1965) can be used to judge further the neighbourhood effect on incidence of adverse birth outcomes. This effect is typically used to measure the loss of

effectiveness by the use of cluster sampling, instead of simple random sampling. The effect is basically the ratio of the actual variance, under the sampling method actually used, to the variance computed under the assumption of simple random sampling. Using this concept, the design effect can be measured by calculating the ratio of the actual variance under the current distribution, to the variance computed under the assumption that the distribution of adverse outcomes is from one random distribution. The design effect is calculated as:

$$DEEF = 1 + ICC * (n - 1), \quad (5.2)$$

where  $n$  is the average neighbourhood sample size, which is calculated by the ratio of the total live births (90500) to the total dissemination areas (1411) in this case. Using this equation, the design effects corresponding to the ICC values in Table 5.6 were calculated. The results are shown in Table 5.7.

<i>Estimator</i>	<i>DEEF for preterm birth</i>	<i>DEEF for LBW</i>
Logit model (SAS GLIMMIX)	2.39	3.04
Logit model (WinBUGS)	2.39	3.12
Probit model (WinBUGS)	1.86	2.19

Table 5.7: Design effect

The design effects can be interpreted such that the variances from the current LBW (or preterm) birth distribution were 1.86-2.39 (or 2.19-3.12) times bigger than they would be if the LBW (or preterm) births were from a random distribution. The identified spatial variances are statistically significant both by the test of the Bayesian modeling or by the frequentist modeling constructed earlier, indicating non-ignorable area-level impacts.

Since the above ICC estimators are based on the assumption that adverse birth outcomes are independently distributed between neighbourhoods, they therefore do not account for any spatial autocorrelation of the neighbourhoods. Hence, a further global Moran's  $I$  test using Equation (4.18) was then carried out to assess if there is a recognizable spatial clustering of adverse birth outcomes.

Instead of directly measuring the Moran's  $I$  score of the raw adverse birth outcome rates, the neighbourhood-level random effect  $v_{0j}$  in Equation (5.1) was used for testing global spatial autocorrelation. The  $v_{0j}$  values obtained by the *logit* model using a Bayesian approach were used for the test. The results are shown in Table 5.8.

<i>Adverse birth outcome</i>	<i>Moran's I value</i>	<i>Z score</i>	<i>P value</i>
Preterm birth	0.88	135.1	< 0.00001
LBW	0.686	105	< 0.00001

Table 5.8: Global Moran's  $I$  for the random effects of LBW and preterm births



Table 5.8 shows that the spatial distributions of the random effects for preterm births and LBW births after controlling personal-level variables had strong spatial clustering effects (0.88 for preterm births and 0.686 for LBW). Both of them are statistically significant. This result suggests that some potential neighbourhood-level risks may exist in the study region. Thus, through the above tests for Hypothesis 1, the null hypothesis  $H_{I0}$  is rejected with high confidence, and the alternative hypothesis  $H_{I1}$ , namely that *Significant spatial differences of adverse birth outcomes exist among different neighbourhoods, is supported*. Significant spatial differences and spatial clustering of adverse birth outcomes are evident at the neighbourhood level in the study region. Although the proportions of the spatial variation compared to the total variation are relatively small, it is necessary to investigate further the spatial impacts of adverse birth outcomes given their statistical significance and their strong spatial autocorrelations.

### 5.3.3 Hypothesis 2: Compositional vs. Contextual Factors

Given the confirmation of significant neighbourhood variations of adverse birth outcomes, the second hypothesis test was conducted to identify further whether the observed neighbourhood variations were a result of the composition of personal risk factors at each neighbourhood, or whether they resulted from neighbourhood characteristics or contextual risks. If spatial variations become ignorable and no significant spatial dependencies remain after controlling personal-level risks, it can be concluded that the spatial variations of adverse birth outcomes were due to compositional impacts of personal risks. On the other hand, if associations between some neighbourhood risk factors and adverse birth outcomes can be established after controlling for personal risks, then it can be said that contextual determinants existed and contributed to the spatial inequalities of adverse birth outcomes.

#### 5.3.3.i Personal and community-level risk identification

Before this test was conducted, various personal and community-level risks known to be associated with adverse birth outcomes in the study region had to be identified. Based on the personal and neighbourhood-level data obtained, LBW and preterm birth outcomes were regressed to potential personal and community-level risks separately by considering only one potential risk factor at a time in order to determine which ones were statistically associated with adverse birth outcomes. All potential personal-level explanatory variables described in Section 5.2.1 were fitted into Equation (4.5) separately to test the associations between personal-level risks and adverse birth outcomes.

Neighbourhood-level potential risk factors fitted into Equation (4.6) included neighbourhood ethnic composition (percentage of immigrants, percentage of visible minorities, percentage of aboriginal people, percentage of people whose official language is not English), neighbourhood living

conditions (average dwelling values, average person per room, average person per bedroom, rent rate, average number of house repairs, residential instability), neighbourhood socio-economic status (average household incomes, average total incomes, average family incomes, average employment incomes, low income rate, unemployment rate, low education rate, and average socio-economic status), food supply (average food insecurity, and average insufficient vegetable intake), neighbourhood health conditions (average chronic health conditions, average self-perceived health, and average self-perceived unmet health care), neighbourhood drinking (average heavy drinkers, and average regular drinkers), neighbourhood smoking behaviour (average daily smoking rate, and average rate of smoking inside home), psycho-social stability of neighbourhood (average sense of community, average self-perceived stress, and average emotional unhappiness), and average neighbourhood physical inactivity.

Statistical significance was tested at the 95% confidence level. Since the procedure in these tests is regarded as exploratory rather than confirmatory, no further correction for multiple tests was conducted. The results for preterm birth and LBW analysis are listed in Tables 5.9 and 5.10 respectively (refer to Appendix A for the full description of variable names).

<i>level of Risk</i>	<i>Effect</i>	<i>Estimate</i>	<i>Standard Error</i>	<i>t Value</i>	<i>Pr &gt;  t </i>
Personal Risks	<b>HEALTHCLNG</b>	<b>1.2826</b>	<b>0.113</b>	<b>11.35</b>	<b>&lt;.0001</b>
	<b>INFECTIONS</b>	<b>0.3841</b>	<b>0.1227</b>	<b>3.13</b>	<b>0.0017</b>
	<b>LANGNOTEN</b>	<b>-0.1372</b>	<b>0.06602</b>	<b>-2.08</b>	<b>0.0377</b>
	<b>DRUGS</b>	<b>0.8366</b>	<b>0.1374</b>	<b>6.09</b>	<b>&lt;.0001</b>
	<b>GENHLTHCLN</b>	<b>0.3098</b>	<b>0.1424</b>	<b>2.18</b>	<b>0.0296</b>
	<b>SINGLEPARN</b>	<b>0.2489</b>	<b>0.07096</b>	<b>3.51</b>	<b>0.0005</b>
	<b>NOSOCIOSPT</b>	<b>0.3168</b>	<b>0.08141</b>	<b>3.89</b>	<b>&lt;.0001</b>
	<b>FINADIFC</b>	<b>0.3673</b>	<b>0.07555</b>	<b>4.86</b>	<b>&lt;.0001</b>
	<b>NOPRENCARE</b>	<b>0.9853</b>	<b>0.1142</b>	<b>8.63</b>	<b>&lt;.0001</b>
	<b>MENTALMOM</b>	<b>0.5919</b>	<b>0.2934</b>	<b>2.02</b>	<b>0.0437</b>
	<b>MARIDISTR</b>	<b>0.5501</b>	<b>0.1908</b>	<b>2.88</b>	<b>0.0039</b>
	<b>FAMILYVIO</b>	<b>0.4459</b>	<b>0.2214</b>	<b>2.01</b>	<b>0.044</b>
	<b>STRSDELIVY</b>	<b>1.5638</b>	<b>0.04795</b>	<b>32.62</b>	<b>&lt;.0001</b>
	<b>SMOKING</b>	<b>0.3668</b>	<b>0.07667</b>	<b>4.78</b>	<b>&lt;.0001</b>
	<b>UNDERNUTRI</b>	<b>1.1981</b>	<b>0.415</b>	<b>2.89</b>	<b>0.0039</b>
	SCHIZMOM	0.403	0.2084	1.93	0.0532
	SCHIZDAD	-0.3911	0.7238	-0.54	0.5889
	MENTALDAD	0.8189	0.6206	1.32	0.187
	LOWEDU	-0.174	0.113	-1.54	0.1236
	NOPRECLASS	-0.05635	0.0761	-0.74	0.459
Neighbourhood	<b>NO_CITIZEN</b>	<b>0.6771</b>	<b>0.3068</b>	<b>2.21</b>	<b>0.0273</b>

Risks	IMMIGRANTS	0.4344	0.1631	2.66	0.0077
	VIS_MINO	0.3909	0.1508	2.59	0.0096
	ABORIGINAL	1.3795	0.6167	2.24	0.0253
	NONOFFLANG	0.377	0.1309	2.88	0.004
	TOT_INC	-0.1073	0.02002	-5.36	<.0001
	EMPLOY_INC	-0.0897	0.02016	-4.45	<.0001
	FAMILY_INC	-0.1119	0.02021	-5.54	<.0001
	HOUSE_INC	-0.1182	0.02002	-5.9	<.0001
	LOW_INCOME	0.7094	0.1476	4.81	<.0001
	RENT_RATE	0.3111	0.06831	4.55	<.0001
	MAJHOSREPA	1.1173	0.3035	3.68	0.0002
	DWL_VAL	-0.1256	0.0205	-6.13	<.0001
	UNEMPLOY	0.01278	0.003713	3.44	0.0006
	LOW_EDU	0.3512	0.08639	4.07	<.0001
	LOW_SES	0.9982	0.1825	5.47	<.0001
	MOVERS	0.1373	0.09629	1.43	0.1538
	PERROOM	-0.06027	0.05939	-1.01	0.3102
	PERBEDROOM	-0.2861	0.1606	-1.78	0.0749
	CHC	-0.8607	0.2201	-3.91	<.0001
	EU	0.7139	0.9313	0.77	0.4433
	FI	1.3091	0.2679	4.89	<.0001
	HD	-0.5911	0.2661	-2.22	0.0264
	IV	0.8914	0.2283	3.90	<.0001
	SPUH	0.5225	0.3723	1.40	0.1605
	RD	-0.3998	0.2320	-1.72	0.0848
	SBC	0.1785	0.1209	1.48	0.1397
	DS	0.02906	0.01596	1.82	0.0687
	PI	0.2563	0.2104	1.22	0.2232
	SIH	0.7339	0.4180	1.76	0.0792
	SPH	-0.1159	0.04557	-2.54	0.0110
	SPS	-0.1150	0.1271	-0.90	0.3656

Table 5.9: Regression results for preterm birth risks

<i>level of Risk</i>	<i>Effect</i>	<i>Estimate</i>	<i>Standard Error</i>	<i>t Value</i>	<i>Pr &gt;  t </i>
Personal risks	DRUGS	1.0141	0.1626	6.24	<.0001
	HEALTHCLNG	0.9125	0.1495	6.1	<.0001
	SINGLEPARN	0.3772	0.0848	4.45	<.0001
	NOSOCIOSPT	0.4749	0.09634	4.93	<.0001
	FINADIFC	0.4281	0.08969	4.77	<.0001
	NOPRENCARE	0.3967	0.1522	2.61	0.0092

	MARIDISTR	0.5047	0.2427	2.08	0.0376
	LOWEDU	0.427	0.1195	3.57	0.0004
	FAMILYVIO	0.7986	0.257	3.11	0.0019
	SMOKING	0.8682	0.08469	10.25	<.0001
	UNDERNUTRI	2.1305	0.4041	5.27	<.0001
	STRSDELIVY	0.888	0.06207	14.31	<.0001
	INFECTIONS	0.2615	0.1545	1.69	0.0906
	LANGNOTEN	0.03598	0.08053	0.45	0.6551
	GENHLTHCLN	-0.0896	0.2021	-0.44	0.6575
	SCHIZMOM	-0.1321	0.2999	-0.44	0.6597
	SCHIZDAD	-0.8601	1.0766	-0.8	0.4244
	MENTALMOM	0.6682	0.3598	1.86	0.0633
	MENTALDAD	0.7602	0.7808	0.97	0.3302
	NOPRECLASS	0.02586	0.09281	0.28	0.7805
Neighbourhood risks	NO_CITIZEN	1.4187	0.3702	3.83	0.0001
	IMMIGRANTS	0.7153	0.2029	3.53	0.0004
	VIS_MINO	0.8408	0.1835	4.58	<.0001
	ABORIGINAL	1.8863	0.7711	2.45	0.0144
	NONOFFLANG	0.7042	0.1613	4.37	<.0001
	DWL_VAL	-0.1076	0.02618	-4.11	<.0001
	PERROOM	-0.3394	0.07376	-4.6	<.0001
	PERBEDROOM	-1.0393	0.1974	-5.27	<.0001
	RENT_RATE	0.4897	0.08367	5.85	<.0001
	MOVERS	0.4205	0.1216	3.46	0.0005
	HOUSE_INC	-0.1201	0.02568	-4.68	<.0001
	TOT_INC	-0.1192	0.02569	-4.64	<.0001
	EMPLOY_INC	-0.1064	0.02589	-4.11	<.0001
	FAMILY_INC	-0.128	0.02594	-4.93	<.0001
	LOW_INCOME	0.8919	0.1808	4.93	<.0001
	UNEMPLOY	0.0161	0.00461	3.49	0.0005
	LOW_EDU	0.379	0.11	3.44	0.0006
	LOW_SES	1.1632	0.2293	5.07	<.0001
	MAJHOSREPA	0.3702	0.3897	0.95	0.3421
	CHC	-0.3999	0.2804	-1.43	0.1537
	EU	3.6438	1.1323	3.22	0.0013
	FI	1.6067	0.3329	4.83	<.0001
	HD	-1.0104	0.3388	-2.98	0.0029
	IV	1.1307	0.2870	3.94	<.0001
	PI	0.5241	0.2669	1.96	0.049
	RD	-0.7744	0.2927	-2.65	0.0082
	SBC	0.3407	0.1506	2.26	0.0237

	DS	0.03775	0.02009	1.88	0.0603
	SIH	-0.2609	0.5316	-0.49	0.6237
	<b>SPUH</b>	<b>1.0137</b>	<b>0.4666</b>	<b>2.17</b>	<b>0.0298</b>
	SPH	-0.06915	0.05806	-1.19	0.2337
	SPS	0.1136	0.1612	0.70	0.4809

Table 5.10: Regression results for LBW risks

As shown in Table 5.9, after controlling for teenage mothers, mothers of advanced age, female births, and multiple births, preterm births in the three health regions are positively associated with personal-level risks, including mothers with health challenges, infections, alcohol or drug abuse during pregnancy, family genetic history of health challenges, single parent, no social support, financial difficulty, no prenatal care, mentally challenged mother, marital distress, family violence, smoking during pregnancy, high stress related to delivery, and maternal under-nutrition, and negatively associated with a minority first language (language not English). Preterm births are also positively associated with neighbourhood-level risks, including ethnic and language minority, low socio-economic status (low income, low education and unemployment), poor living conditions, insufficient food and vegetable supply, and negatively associated with average regular drinking, self-perceived poor health conditions, and average chronic health conditions. All associations are statistically significant at or beyond the 95% confidence level. Preterm births are not associated with evidence of schizophrenia of mother or father, mentally challenged father, low maternal education, and not attending prenatal classes.

Table 5.10 shows that, after controlling for teenage mothers, mothers of advanced age, babies' sex, multiple births, and preterm births, LBW births are positively associated with personal-level risks including health challenges, drug or alcohol abuse during pregnancy, single parent, no social support, financial difficulty, no prenatal care, marital distress, low education, family violence, smoking, maternal under-nutrition, and high stress related to delivery. LBW births are also positively associated with neighbourhood-level variables including ethnic and language minority, poor living conditions, low socio-economic status (low income, low education, and unemployment), insufficient food and vegetable supply, lack of sense of community, average emotional unhappiness, and physical inactivity, and are negatively associated with average heavy and regular drinking. Again all associations are statistically significant at or beyond the 95% confidence level. Personal factors, such as infections, non-English speaking mothers, family history of genetic health challenges, evidence of schizophrenia of mother or father, mentally challenged mother or father, and not attending prenatal classes are not associated with LBW.

The above results show that adverse birth outcomes are associated with numerous and complex

personal and neighbourhood level risk factors including health, behavioural, cultural, socio-economic, psycho-social, and environmental risks. Most of the identified associations are consistent with the literature discussed in Chapter 2. It can also be observed that although they share many of the same risk factors, the risks of having preterm births and LBW are different, in terms of both risk factors and magnitude. Maternal infection is associated with preterm birth but not associated with LBW. This is consistent with the literature which suggests that infection plays a more important role in the onset of preterm births than in IUGR births. All genetic health-related risks are not related to LBW, indicating a lack of support for a hypothesis of direct family genetic impacts on LBW. Two genetic-related variables (family history of genetic health challenges and mentally challenged mother) are associated with preterm births, suggesting some degree of genetic impact. However, these associations may only be due to the adverse situations and consequences associated with genetic health problems, such as low socio-economic status or adverse health behaviours.

The failure to attend prenatal classes seems to have nothing to do with adverse birth outcomes. Most prenatal health education through prenatal classes focuses mainly on suggestions of general care plans, delivery preparation, infant care, and breast feeding. There is only some very brief content regarding risk factors of adverse birth outcomes. While knowledge gained through these prenatal classes may be helpful for maternal health and baby growth in general, they may not be helpful for preventing adverse birth outcomes. This insignificant association may also be due to the attendance of prenatal classes at a late stage of pregnancy. It is possible that education-based prevention procedures could be effective if taken during an early stage of pregnancy and more knowledge of adverse birth outcome prevention is provided.

Although the percentage of minority ethnic groups is positively related to both preterm births and LBW at the neighbourhood level, personal-level non-official-language-speaking shows a negative association with preterm births, and non-association with LBW. This might suggest that while within-group integration of minority ethnic groups may help to reduce preterm births for mothers within minority groups, neighbourhoods that have a high percentage of minority ethnic populations may create an environment that increases the overall risks of having adverse birth outcomes. The negative associations of average heavy drinking, self-perceived health, and chronic health conditions with preterm births are somewhat counterintuitive. The negative associations of average regular and heavy drinking with LBW are counterintuitive as well. While personal-level health conditions and drug and alcohol abuse did show positive impacts on adverse birth outcomes, community-level chronic health conditions and alcohol use show the opposite impacts on preterm births. Since neighbourhoods having a high percentage of residents with chronic conditions will normally receive more care and a high percentage of drinking will indicate more neighbourhood

social activities, increased neighbourhood integration may explain the decreased adverse birth outcomes. These phenomena are further analyzed and discussed in Section 5.3.3.

The associations with living condition variables also suggest that preterm births may be more associated with localized environmental pollution (e.g. environmental damage caused by major house repairs), and LBW births may be more associated with overcrowded living environments (such as high rates of persons per room and persons per bedroom). Neighbourhood socio-economic status is the most consistent neighbourhood indicator associated with adverse birth outcomes. All income, education, and employment variables show consistent associations with adverse birth outcomes. Hence, further analyses are needed to identify statistically the pathways of social and environmental impacts on adverse birth outcomes.

#### 5.3.3.ii Factor analysis

The many identified personal-level risks of both preterm births and LBW represent several different aspects of personal-level characteristics. Some of them are highly correlated, which constrains the ability to create an integrated model to depict the overall impacts of personal-level risks on adverse birth outcomes. This is important not only for the test of contextual effects, but also for the later mediational effect tests. To eliminate the problems caused by the correlated personal-level risk factors, the identified personal-level risks in Tables 5.9 and 5.10 were used to conduct a factor analysis using the methods described in Chapter 4.4.3. However, maternal under-nutrition was excluded due to its lack of representation at the CDA level. Compared to the number of CDAs (1411) in the study region, the proportion of mal-nourished mothers is only 0.05% of all the live births (49 out of 90485). Hence, this variable therefore cannot effectively represent neighbourhood variations of maternal under-nutrition conditions.

A principal component analysis was first conducted on the identified personal-level risks of preterm births and LBW respectively. The optimal number of factors was determined based on the criteria that each factor should have an eigenvalue greater than the average of the initial communality estimates, and all of the common variance (defined by the sum of communality estimates) should be explained by extracted factors. This ensures that the common variance represented by the original personal risk variables can be completely accounted for by the extracted factors without losing any explanatory power. The identified principal factors were then rotated using the varimax orthogonal rotation method to determine the best combination of the personal-level risks to represent different aspects of personal characteristics. The factor patterns for preterm births and LBW are listed in Tables 5.11 and 5.12 respectively, where only large factor loadings ( $\geq 0.3$ ) are displayed so that the major contribution of risk variables to each factor can be presented clearly. Although factor loadings

less than 0.3 are not displayed in the tables, they were still used to estimate the factor scores even though their impacts on corresponding factors were small. Considered as explanatory variables, the factors were centred with 0 means and are approximately orthogonal.

<i>Variable name</i>	<i>Factor 1</i>	<i>Factor 2</i>	<i>Factor 3</i>	<i>Factor 4</i>	<i>Factor 5</i>
HEALTHCLNG	.	.	.	0.5040	.
INFECTIONS	.	.	0.4097	.	.
DRUGS	0.3754	.	0.6059	.	.
GENHLTHCLN	.	.	.	.	0.6849
SINGLEPARN	0.6172	0.411	0.3843	.	.
NOSOCIOSPT	0.7654	.	.	.	.
FINADIFC	0.7206	.	0.3291	.	.
NOPRENCARE	0.5085	.	0.4062	.	.
MENTALMOM	0.3204	.	.	.	0.6163
MARIDISTR	.	0.7859	.	.	.
FAMILYVIO	.	0.6728	0.3544	.	.
STRSDELIVY	.	.	.	0.4839	.
SMOKING	.	0.3369	0.5097	.	.
	<b>SES</b>	<b>Psycho-Social</b>	<b>Behavioural</b>	<b>Health</b>	<b>Genetic</b>

Table 5.11: Factor pattern for preterm births

<i>Variable name</i>	<i>Factor 1</i>	<i>Factor2</i>	<i>Factor 3</i>	<i>Factor 4</i>
HEALTHCLNG	.	.	.	0.5425
DRUGS	0.5359	0.3740	.	.
SINGLEPARN	0.4140	0.5683	0.4294	.
NOSOCIOSPT	.	0.7418	.	.
FINADIFC	0.5140	0.6382	.	.
NOPRENCARE	0.4552	0.4824	.	.
MENTALMOM	0.5989	.	.	.
MARIDISTR	.	.	0.7662	.
LOWEDU	0.6748	.	0.3080	.
FAMILYVIO	0.3984	.	0.6560	.
STRSDELIVY	.	.	0.3217	0.4203
SMOKING	0.6439	.	0.3642	.
	<b>Behavioural</b>	<b>SES</b>	<b>Psycho-Social</b>	<b>Health</b>

Table 5.12: Factor pattern for LBW

In Table 5.11, Factor 1 is mainly composed of single parent, no social support, financial



difficulty, and no prenatal care, all of which represent a lack of socio-economic resources. Thus, Factor 1 can be called a SES factor. Factor 2 is mainly composed of single parent, marital distress, family violence, and smoking. These variables have a common characteristic of representing high stress or depression of the mother. Factor 2 can therefore be called a psycho-social factor. Factor 3 is mainly composed of drug use, no prenatal care, infections, and smoking, and somewhat related to single parent, family violence, and financial difficulties. The common character among these variables is that all these may represent careless behaviours of the mother. Hence, this factor can be called a behavioural factor. Factor 4 is mainly composed of mother's health challenges and stress related to delivery. This is straightforward to interpret as a health factor. Factor 5 is composed of family history of genetic health challenges and mentally challenged mother and is considered to be a genetic factor.

After these factors were constructed, they were fitted into the model to test their associations with preterm births. The model is constructed as:

*Adverse birth outcomes* ~ *binary* ( $p_{ij}$ )

$$\text{Level 1 (personal): } \text{logit}(p_{ij}) = \beta_{0j} + \beta_1 \text{AGE19}_{ij} + \beta_2 \text{AGE36}_{ij} + \beta_3 \text{FEMALE}_{ij} + \beta_4 \text{MLTIBIRTH}_{ij} \\ + \beta_5 \text{factor1}_{ij} + \beta_6 \text{factor2}_{ij} + \beta_7 \text{factor3}_{ij} + \beta_8 \text{factor4}_{ij} + \beta_9 \text{factor5}_{ij}$$

$$\text{Level 2 (neighbourhood): } \beta_{0j} = \gamma_{00} + v_{0j}. \quad (5.3)$$

The computed result for the fixed effects is presented in Table 5.13. The results show that Factor 5 was not statistically associated with preterm births measured at the 95% confidence level, suggesting a lack of association between preterm births and mother's family genetic history. This confirms the early assumption that preterm births are not caused directly by genetic differences, but by the consequences or difficulties brought about by genetic health problems.

<i>Effect</i>	<i>Estimate</i>	<i>Standard Error</i>	<i>t Value</i>	<i>Pr &gt;  t </i>
<b>Intercept</b>	<b>-2.7852</b>	<b>0.0235</b>	<b>-118.53</b>	<b>&lt;.0001</b>
<b>AGE19</b>	<b>0.1999</b>	<b>0.0682</b>	<b>2.93</b>	<b>0.0034</b>
<b>AGE36</b>	<b>0.1991</b>	<b>0.0433</b>	<b>4.6</b>	<b>&lt;.0001</b>
<b>FEMALE</b>	<b>-0.1308</b>	<b>0.0286</b>	<b>-4.57</b>	<b>&lt;.0001</b>
<b>MULTIBIRTH</b>	<b>2.9129</b>	<b>0.0427</b>	<b>68.15</b>	<b>&lt;.0001</b>
<b>Factor1</b>	<b>0.0993</b>	<b>0.0167</b>	<b>5.96</b>	<b>&lt;.0001</b>
<b>Factor2</b>	<b>0.0755</b>	<b>0.0141</b>	<b>5.37</b>	<b>&lt;.0001</b>
<b>Factor3</b>	<b>0.1300</b>	<b>0.0177</b>	<b>7.34</b>	<b>&lt;.0001</b>
<b>Factor4</b>	<b>0.4416</b>	<b>0.0171</b>	<b>25.83</b>	<b>&lt;.0001</b>
Factor5	0.0237	0.0195	1.22	0.2239

Table 5.13: The association of personal-level risk factors with preterm births

Similarly, for LBW (Table 5.12), Factor 1 is a behavioural factor, Factor 2 is a SES factor, Factor 3 is a psycho-social factor, and Factor 4 is a health factor. A similar test as for Equation (5.3) was conducted. The results (Table 5.14) show that all of the four constructed factors were positively associated with LBW.

<i>Effect</i>	<i>Estimate</i>	<i>Standard Error</i>	<i>t Value</i>	<i>Pr &gt;  t </i>
<b>Intercept</b>	<b>-4.0906</b>	<b>0.0351</b>	<b>-116.58</b>	<b>&lt;.0001</b>
<b>AGE19</b>	<b>0.2736</b>	<b>0.0824</b>	<b>3.32</b>	<b>0.0009</b>
<b>AGE36</b>	<b>0.1644</b>	<b>0.0558</b>	<b>2.94</b>	<b>0.0032</b>
<b>FEMALE</b>	<b>0.2341</b>	<b>0.0367</b>	<b>6.39</b>	<b>&lt;.0001</b>
<b>MULTIBIRTH</b>	<b>2.0394</b>	<b>0.0572</b>	<b>35.65</b>	<b>&lt;.0001</b>
<b>PRETERM</b>	<b>3.5432</b>	<b>0.0375</b>	<b>94.48</b>	<b>&lt;.0001</b>
<b>Factor1</b>	<b>0.1628</b>	<b>0.0206</b>	<b>7.91</b>	<b>&lt;.0001</b>
<b>Factor2</b>	<b>0.0817</b>	<b>0.0239</b>	<b>3.42</b>	<b>0.0006</b>
<b>Factor3</b>	<b>0.1298</b>	<b>0.0170</b>	<b>7.63</b>	<b>&lt;.0001</b>
<b>Factor4</b>	<b>0.2434</b>	<b>0.0227</b>	<b>10.73</b>	<b>&lt;.0001</b>

Table 5.14: The association of personal-level risk factors with LBW

The first four factors for preterm births and all of the four factors for LBW were used instead of the original personal-level risk variables for the following analyses. These factors are independent to each other and cover most of the variations of identified personal-level risks of adverse birth outcomes. They can therefore be used as common personal-level risks without losing any generalizability.

### 5.3.3.iii Contextual effect testing

Given all of the identified personal-level risks, the contextual effects of neighbourhood-level risks were tested. A model is constructed as:

*Adverse birth outcomes* ~ *binary* ( $p_{ij}$ )

$$\text{Level 1 (personal): } \logit(p_{ij}) = \beta_{0j} + \beta_1 \text{AGE19}_{ij} + \beta_2 \text{AGE36}_{ij} + \beta_3 \text{FEMALE}_{ij} + \beta_4 \text{MLTIBIRTH}_{ij} \\ + \beta_5 \text{factor1}_{ij} + \beta_6 \text{factor2}_{ij} + \beta_7 \text{factor3}_{ij} + \beta_8 \text{factor4}_{ij}$$

$$\text{Level 2 (neighbourhood): } \beta_{0j} = \gamma_{00} + \gamma_{01} \text{NB\_RISK}_j + v_{0j}. \quad (5.4)$$

Using this model, neighbourhood-level risks identified in Tables 5.9 and 5.10 were fitted separately one at a time. The results are shown in Tables 5.15 and 5.16. It can be observed in Table 5.15 that, after controlling for personal-level risks, most of the neighbourhood low socio-economic status variables (except unemployment), physical living conditions (high rent rate, and major house repair), food and vegetable insecurity, were still positively associated with preterm births, indicating

some direct impacts of these neighbourhood risk factors on the foetus. For LBW (Table 5.16), neighbourhood ethnic minority, low socio-economic status, physical living conditions, food and vegetable insecurity, emotionally unhappiness, and sense of not belonging to local community were positively associated with LBW after controlling personal-level risks. These results support the alternative of Hypothesis 2,  $H2_1$ , namely that contextual risks contribute to spatial inequalities of adverse birth outcomes after controlling for personal compositional risks. The null hypothesis,  $H2_0$ , is therefore rejected.

<i>Effect</i>	<i>Estimate</i>	<i>Standard Error</i>	<i>t Value</i>	<i>Pr &gt;  t </i>
NO_CITIZEN	0.2354	0.3051	0.77	0.4402
IMMIGRANTS	0.2746	0.1616	1.70	0.0893
VIS_MINO	0.1861	0.1498	81044	0.2141
ABORIGINAL	0.7342	0.6171	1.19	0.2341
NONOFFLANG	0.2001	0.1302	1.54	0.1243
TOT_INC	<b>-0.0767</b>	<b>0.0199</b>	<b>-3.86</b>	<b>0.0001</b>
EMPLOY_INC	<b>-0.0618</b>	<b>0.0200</b>	<b>-3.09</b>	<b>0.0020</b>
FAMILY_INC	<b>-0.0784</b>	<b>0.0201</b>	<b>-3.90</b>	<b>&lt;.0001</b>
HOUSE_INC	<b>-0.0839</b>	<b>0.0200</b>	<b>-4.20</b>	<b>&lt;.0001</b>
LOW_INCOME	<b>0.4234</b>	<b>0.1493</b>	<b>2.84</b>	<b>0.0046</b>
RENT_RATE	<b>0.1932</b>	<b>0.0690</b>	<b>2.80</b>	<b>0.0051</b>
MAJHOSREPA	<b>0.8435</b>	<b>0.3018</b>	<b>2.80</b>	<b>0.0052</b>
DWL_VAL	<b>-0.0932</b>	<b>0.0203</b>	<b>-4.59</b>	<b>&lt;.0001</b>
UNEMPLOY	0.0071	0.0037	1.92	0.0553
LOW_EDU	<b>0.2395</b>	<b>0.0861</b>	<b>2.78</b>	<b>0.0054</b>
LOW_SES	<b>0.6394</b>	<b>0.1839</b>	<b>3.48</b>	<b>0.0005</b>
SPH	-0.0651	0.0449	-1.45	0.1475
CHC	<b>-0.6594</b>	<b>0.2179</b>	<b>-3.03</b>	<b>0.0025</b>
HD	-0.4237	0.2633	-1.61	0.1075
FI	<b>0.9127</b>	<b>0.2685</b>	<b>3.40</b>	<b>0.0007</b>
IV	<b>0.6393</b>	<b>0.2268</b>	<b>2.82</b>	<b>0.0048</b>

Table 5.15: Contextual effects of neighbourhood variables on preterm births

The magnitude of most of the neighbourhood risks were reduced after controlling personal-level risks and some of the neighbourhood risks reduced to insignificance, suggesting that neighbourhood risks are partially or totally mediated by personal risks. Thus, both compositional and contextual factors play important roles in determining adverse birth outcomes. These results show that instead of investigating solely personal-level risk factors, a direct intervention on the neighbourhood-level risks

identified in Tables 5.15 and 5.16 is necessary to reduce effectively adverse birth outcomes.

<i>Effect</i>	<i>Estimate</i>	<i>Standard Error</i>	<i>t Value</i>	<i>Pr &gt;  t </i>
<b>NO_CITIZEN</b>	<b>1.0674</b>	<b>0.3711</b>	<b>2.88</b>	<b>0.0040</b>
<b>IMMIGRANTS</b>	<b>0.6152</b>	<b>0.2020</b>	<b>3.05</b>	<b>0.0023</b>
<b>VIS_MINO</b>	<b>0.6850</b>	<b>0.1837</b>	<b>3.73</b>	<b>0.0002</b>
ABORIGINAL	1.2589	0.7722	1.63	0.1030
<b>NONOFFLANG</b>	<b>0.5691</b>	<b>0.1612</b>	<b>3.53</b>	<b>0.0004</b>
<b>DWL_VAL</b>	<b>-0.0738</b>	<b>0.0261</b>	<b>-2.83</b>	<b>0.0047</b>
<b>PERROOM</b>	<b>-0.2946</b>	<b>0.0736</b>	<b>-4.01</b>	<b>&lt;.0001</b>
<b>PERBEDROOM</b>	<b>-0.8411</b>	<b>0.1977</b>	<b>-4.25</b>	<b>&lt;.0001</b>
<b>RENT_RATE</b>	<b>0.3877</b>	<b>0.0850</b>	<b>4.56</b>	<b>&lt;.0001</b>
<b>MOVERS</b>	<b>0.3658</b>	<b>0.1200</b>	<b>3.05</b>	<b>0.0023</b>
<b>HOUSE_INC</b>	<b>-0.0850</b>	<b>0.0257</b>	<b>-3.31</b>	<b>0.0009</b>
<b>TOT_INC</b>	<b>-0.0864</b>	<b>0.0257</b>	<b>-3.37</b>	<b>0.0008</b>
<b>EMPLOY_INC</b>	<b>-0.0756</b>	<b>0.0258</b>	<b>-2.93</b>	<b>0.0034</b>
<b>FAMILY_INC</b>	<b>-0.0932</b>	<b>0.0259</b>	<b>-3.59</b>	<b>0.0003</b>
<b>LOW_INCOME</b>	<b>0.6237</b>	<b>0.1841</b>	<b>3.39</b>	<b>0.0007</b>
<b>UNEMPLOY</b>	<b>0.0109</b>	<b>0.0046</b>	<b>2.35</b>	<b>0.0188</b>
<b>LOW_EDU</b>	<b>0.2613</b>	<b>0.1101</b>	<b>2.37</b>	<b>0.0176</b>
<b>LOW_SES</b>	<b>0.8088</b>	<b>0.2323</b>	<b>3.48</b>	<b>0.0005</b>
<b>EU</b>	<b>3.4954</b>	<b>1.7399</b>	<b>2.01</b>	<b>0.0445</b>
<b>FI</b>	<b>0.6750</b>	<b>0.2541</b>	<b>2.66</b>	<b>0.0079</b>
<b>HD</b>	<b>-1.5835</b>	<b>0.4575</b>	<b>-3.46</b>	<b>0.0005</b>
<b>IV</b>	<b>1.1909</b>	<b>0.3716</b>	<b>3.20</b>	<b>0.0014</b>
PI	0.4819	0.3625	1.33	0.1837
<b>RD</b>	<b>-1.0780</b>	<b>0.4165</b>	<b>-2.59</b>	<b>0.0096</b>
<b>SBC</b>	<b>0.5468</b>	<b>0.2201</b>	<b>2.48</b>	<b>0.0130</b>
SPUH	0.7669	0.4432	1.73	0.0836

Table 5.16: Contextual effects of neighbourhood variables on LBW

For compositional impacts, intervention on personal-level risk factors may be effective. However, it is also necessary to identify the pathways through which neighbourhood-level risks are mediated or modified by personal-level risks so that a more effective and targeted intervention can be derived, instead of a universal intervention on all population and personal risk factors. In the following section, a mediational analysis is conducted to identify the pathways of neighbourhood-level risks and to test the third and fourth research hypotheses.

### 5.3.4 Mediation analysis on pathways of neighbourhood determinants of adverse birth outcomes

Using the methods discussed in Chapter 4.2.2, mediational effects of the neighbourhood-level risks identified in Tables 5.9 and 5.10 were tested using frequentist binary-outcome multilevel logistic modeling (Equations 4.8 through 4.10). The Statistical Analysis System (SAS) code for the mediational testing for preterm births and LBW is provided in Appendix B. The results are listed in Appendices C and D for preterm births and LBW respectively. Tables in Appendices C and D show that the total effects,  $c$ , are approximately equal to the indirect effect plus the direct effects,  $ab+c'$ . This proves that the proposed procedures discussed in Chapter 4.2.2 are reliable to use. Based on these results, Hypotheses 3 and 4 were tested as follows.

#### 5.3.4.i Hypothesis 3: Psycho-social vs. Material Influences

As identified earlier, almost all neighbourhood-level socio-economic related risks investigated were found to have some impact on adverse birth outcomes. Many of them remained associated with adverse birth outcomes after controlling for personal-level risk factors (Tables 5.15 and 5.16), suggesting a direct impact of these factors on the foetus. This finding favours a direct purely material model in that adverse birth outcome variations of different neighbourhood socio-economic groups are determined by their direct exposure to risks or physical hazards in their neighbourhood environment. The significant association after controlling for personal risks indicates that these neighbourhood risks are socially structured and the individual may have little control over these risks. For preterm births, the contextual neighbourhood risks include unhealthy food supply, no home ownership, uncomfortable living conditions due to the need for major house repairs, and other unidentified neighbourhood risks. For LBW births, the risks may include overcrowded living conditions, no home ownership, residential instability, and an unhealthy food supply. This is consistent with the second alternative of Hypothesis 3,  $H3_2$ .

In addition to their direct environmental risk exposures, various socio-economic factors also exert their impacts on adverse birth outcomes through the mothers' personal-level risks. It is evident in Appendices C and D that neighbourhood low socio-economic status (including income, education, employment, rent rate, major house repair, overcrowding, and food security) affects mothers' personal-level risks through almost all four pathways. This is consistent with an interpretation that low neighbourhood SES results in low personal SES, and leads to poor health for the new mothers, puts psycho-social stress on them, and consequently induces their health-threatening behaviours during pregnancy, all of which lead to increased risks of LBW and preterm births. The only difference is that, at the neighbourhood level, a higher proportion of low education residents does not

directly lead to a higher proportion of health-challenged mothers. Low education appears to exert its impact on LBW and preterm birth mainly through the other three personal risk factors. This finding supported both of the two alternatives of Hypothesis 3,  $H3_1$  and  $H3_2$ .

Neighbourhoods with a high proportion of major house repairs required may represent old communities with unhygienic living environments and occupied by low SES residents. In addition to its direct environmental impact on preterm births, the risk of major house repairs required is mediated by mothers' health behaviours, SES, and psycho-social risks, indicating the adverse impacts of a poor living environment on personal behaviour, socio-economic status, and psycho-social conditions. It also has a positive moderation effect on the impact of maternal SES to preterm births. This means that a low personal SES for mothers living in a neighbourhood with many houses requiring major repairs and having lack of social support, financial difficulties, and single motherhood, is associated with an elevated risk of preterm births compared with low SES mothers living in a neighbourhood with relatively new and better houses. This implies some accumulated environmental and social disadvantages on preterm births. The neighbourhood rent rate has a negative moderation effect on the impacts of maternal health conditions on LBW and preterm births. This is probably because mothers living in high rent neighbourhoods are relatively young (it is shown by the data that rent rate is negatively associated with mother's age in the study region). Adverse health conditions may affect adverse birth outcomes less for young mothers than for older mothers. Overall, based on the above identified pathways of neighbourhood-level person per room, rent rate, and residential instability to LBW, and neighbourhood-level major house repairs and rent rate to preterm births, the direct material impact of living conditions on mothers' health is less supported than for other pathways.

Neighbourhood food insecurity and average insufficient vegetable intake not only represent different individual life styles among different neighbourhoods, but also represent unhealthy or insufficient food supplies. These two neighbourhood risks are not only associated with low maternal SES and poor maternal health, but also put psycho-social stress on mothers and affect their health behaviours. As they also have direct impacts on adverse birth outcomes, they are important factors that need to be addressed in community-level interventions.

The above evidence supports the idea that socio-economic risks determine mothers' relative social position, and put psychosocial stresses on mothers of lower social status and consequently generate health-threatening behavioural responses during pregnancy. In this regard, strong neighbourhood social capital should have a buffering effect on adverse birth outcomes.

For preterm births, directly measured neighbourhood variables related to social capital, such as the sense of belonging to local communities, average emotional unhappiness, and average self-perceived stress, do not show association with preterm births, although the sense of belonging to local

community variable is barely insignificant at the 5% level. However, some other neighbourhood-level indicators, such as hard drinking, average chronic health conditions, and average self-perceived health may indirectly represent the social capital of neighbourhoods. Heavy drinking and regular drinking may represent social drinking with increased neighbourhood social activities. Neighbourhoods that have higher heavy drinking rates may imply better social integration and stronger social ties than low drinking rate neighbourhoods. It shows in Appendix C that mothers in these former neighbourhoods are healthier, have higher SES, better psycho-social status, and healthier behaviours. The impact of neighbourhoods with heavy drinking is completely mediated by these personal factors, indicating that neighbourhood heavy drinking affects preterm births through personal factors only. While community-level chronic health conditions and self-perceived health may put stresses on the local residents, these conditions may also associate with increased social care. Residents with worse chronic health conditions and low self-perceived health are generally older.

Since these indicators are self-reported variables, people who are concerned more about their health sometimes exaggerate their health conditions and report worse health. In other words, they are more conscious of their health than others. Neighbourhoods with older and more health-conscious residents are generally better-off economically and residents may devote more resources to invest in their health. It is shown in Appendix C that mothers in such neighbourhoods actually have better health, less family violence, less marital distress, and better maternal SES. This indicates that increased social care may help to reduce preterm births through both personal psycho-social and material risks. The results also show contextual impacts of neighbourhood chronic health conditions and self-perceived health on preterm births, implying some direct impacts of social care to the foetus.

For LBW, some directly measured neighbourhood social capital-related variables, such as emotional unhappiness and sense of not belonging to local communities, do show positive impacts on LBW. The neighbourhood-level variable emotional unhappiness is associated with LBW after controlling personal risks, indicating direct neighbourhood psycho-social impacts on LBW. It also indirectly affects LBW through increased personal psycho-social stress, adverse health behaviours and low SES. While average neighbourhood-level emotional unhappiness mainly measures the emotional aspect of social capital, the sense of belonging to local communities may also measure some other aspects of social capital, such as informational and instrumental supports and collective efficacy. A higher sense of belonging to local communities may sometimes not increase the happiness of local residents, but it may increase the collective efficacy of the neighbourhood. Local residents may participate more in collective actions to address common health related issues in their neighbourhood. As shown in Figure 5.18, the impact of sense of belonging to local communities on LBW is completely mediated by personal-level health behaviours, SES and maternal health. This

means that strong neighbourhood social bonds may help to regulate health behaviours, improve maternal health, and reduce SES related risks (such as lack of social support, single parent, financial difficulties, and lack of prenatal care) and eventually reduce LBW. As discussed earlier, neighbourhood-level heavy drinking and regular drinking may represent increased neighbourhood social activities and increased social ties. These two factors also help to reduce the risks of having LBW.

Neighbourhood-level physical inactivity is positively associated with LBW and the association is completely mediated by maternal SES. This means that in neighbourhoods where residents have a less active life style, mothers' SES is generally low. It is more plausible that the low personal SES leads to physical inactivity. Low SES individuals are more sedentary than the socio-economically better-off population due to the lack of resources and incentive to improve their level of physical activity. Scarcity of facilities or places may be the reason reducing the overall physical activity in neighbourhoods with residents of low SES. Thus, attempts at improving neighbourhood-level physical activity may not help to reduce LBW, since the attempts may not help to reduce the negative impact that personal low SES may bring to LBW.

Neighbourhood-level self-perceived unmet health needs is also positively associated with LBW. This variable reflects potential insufficient health services at the local level. In Appendix D, the impact of insufficient health services on LBW is shown to be completely mediated by mothers' low SES, adverse maternal health and psycho-social stress. It is easy to understand that a mother living in a neighbourhood that lacks health services will have poor health, be psycho-socially stressed, and have little prenatal care. However, the relationship of neighbourhood health care and personal SES may also go the other way. As explained by the theories of political economy and the landscape of collective consumption, low SES people possess less power and may not be able to determine the direction of local health care policies and the deployment of resources. Decisions are made partly to favour groups or organizations that possess power, rather than reflecting the public "need" of health services. Although, as noted earlier, health care is universally insured for Canadians, the decisions on health care distributions may result in local variations in health infrastructure, which influences local residents' health care uses. Procedures need to be taken to involve local residents in socio-economically disadvantaged neighbourhoods to participate in decision making to achieve their collective goals of improving local health care services and other health-related conditions together in a complex interactive manner.

Although the major focus of the above analysis is on the downstream impacts of neighbourhood low SES on personal-level risks, it is worth noting that the association between low neighbourhood SES and low personal SES is complex. Low SES neighbourhoods clearly are mainly comprised of



low SES individuals. However, once the concentration of low SES individuals reaches a certain level, it creates an environment that constrains the personal development of individuals who live in such neighbourhoods by constraining their social contacts, job and education opportunities, and marital partner selection. This creates a downward spiral, in which low neighbourhood SES and low personal SES are intertwined and finally affect adverse birth outcomes.

Based on the above discussion, neighbourhood-level SES-related risks exert their impacts on adverse birth outcomes through both the psycho-social and material pathways, although neighbourhood-level psycho-social factors seem to play a more important role in the onset of LBW (or IUGR) than preterm births. Hence, the null hypothesis,  $H3_0$ , is rejected and the two alternative hypotheses,  $H3_1$  and  $H3_2$ , namely that neighbourhood socio-economic risks affect adverse birth outcomes through both mothers' relative positions among neighbourhoods and through associated material conditions, are supported.

#### 5.3.4.ii Hypothesis 4: Personal vs. Cultural Influences

Neighbourhood ethnic composition is another important aspect of neighbourhood-level impacts on adverse birth outcomes. Different neighbourhood-level ethnicity variables, including percentage of non-Canadian citizens, percentage of immigrants, percentage of visible minorities, percentage of aboriginal population, and percentage of non-official-language-speaking population, show positive associations with both LBW and preterm births. As discussed in Chapter 2, these associations are due to either ethnic predispositions of hereditary risks or group level influences such as cultural impacts or socio-economic disadvantages that are associated with minority groups. Since adverse pregnancy outcomes are generally found among racial and ethnic groups, such as African Americans in the United States, Asians in the United Kingdom, Aborigines in Australia, Negroes and Asians in the Netherlands, and new immigrants in Canada, the approach in this thesis treats ethnic groups as a whole, instead of analyzing different ethnic groups separately. This is based on the assumption and the above evidence that different ethnic minority groups will experience similar negative impacts of adverse birth outcomes for their group members. This assumption is valid also because the purpose is to identify the general pathways of ethnicity impacts, rather than to seek different cultural or genetic impacts for different ethnic/racial groups.

Although various neighbourhood-level ethnic compositions were obtained, it is difficult to obtain detailed ethnic/racial information for individual mothers. Hence, the only ethnicity-related variable obtained for the analysis is non-English speaking mother. Other genetic-related personal data include family history of genetic health challenges, mentally challenged parents, and schizophrenia of parents. These genetically related variables may indirectly test the impact of ethnic genetic influences on

adverse birth outcomes.

It was discussed in Section 5.3.3 that individual genetically-related factors do not have direct impacts on adverse birth outcomes. Factors, such as family history of genetic health challenges, mentally challenged parents and schizophrenia of parents, were not associated with LBW (Table 5.10). Although family history of genetic health challenges and mentally challenged mothers were associated with preterm births (Table 5.9), the association disappeared for the derived genetic factor (Factor 5) (Table 5.13), which is composed mainly of these two variables (Table 5.11). Since the variable of mentally challenged mothers showed statistically significant correlations to behavioural risks, the disappeared association implies that it is the associated socio-economic and behavioural disadvantages of genetically-related health problems rather than genetic predispositions that affect adverse birth outcomes. This supports the view indirectly that the genetic composition of ethnic/racial groups may not affect adverse birth outcomes. In addition, non-English speaking mothers have a smaller risk of having preterm births (Table 5.9) than English-speaking mothers and have no risk of having LBW babies (Table 5.10). It was also confirmed that there is no statistically significant modification effect of non-English speaking mothers on the behavioural factors for both LBW and preterm births, although weak but significant negative correlations between non-English speaking mothers and maternal drug uses ( $-0.016$ ,  $p < 0.001$ ) and between non-English speaking mothers and maternal smoking ( $-0.022$ ,  $p < 0.001$ ) exist. Thus, adverse birth outcomes and maternal behavioural differences are not affected by the inherent differences of maternal ethnic status, and neighbourhood ethnic impacts on adverse birth outcomes may be due to other influences.

To test the null hypothesis,  $H4_0$ , that *there is no direct or indirect impact of neighbourhood ethnic composition on adverse birth outcomes or on related health behaviour during pregnancy*, Pearson correlations were calculated among variables representing different ethnic compositions and low neighbourhood SES so that potential confounding factors could be identified. The results, listed in Table 5.17, show that, except for Aboriginal population, other ethnic composition variables are highly correlated. This likely means that immigrants are mostly non-English speaking, and many of them are visible minorities and do not have Canadian citizenship yet. This may also occur because different ethnic minority groups may tend to live together in the same neighbourhoods. The Aboriginal population is more related to low SES than other ethnic groups, which are only moderately associated with low SES. Since Aboriginal population is not associated with other ethnic status, it is necessary to separate it from other ethnic groups for the following analysis and hypothesis testing.

	<i>NO_CITIZEN</i>	<i>IMMIGRANTS</i>	<i>VIS_MINO</i>	<i>NONOFFLANG</i>	<i>ABORIGINAL</i>	<i>LOW_SES</i>
<i>NO_CITIZEN</i>	1.00	0.6	0.54	0.55	-	0.28
<i>IMMIGRANTS</i>	0.6	1.00	0.67	0.8	-0.05	0.09
<i>VIS_MINO</i>	0.54	0.67	1.00	0.67	0.07	0.24
<i>NONOFFLANG</i>	0.55	0.8	0.67	1.00	-	0.24
<i>ABORIGINAL</i>	-	-0.05	0.07	-	1.00	0.32
<i>LOW_SES</i>	0.28	0.09	0.24	0.24	0.32	1.00

Table 5.17: Pearson Correlation Coefficients of neighbourhood ethnic variables and SES (variable names are defined in Appendix A)

If the null hypothesis holds true, the impacts of neighbourhood ethnic composition should disappear after controlling for confounding neighbourhood-level factors, especially neighbourhood-level SES. Based on these considerations, a model was constructed as follows:

*Adverse birth outcomes ~ binary ( $p_{ij}$ )*

$$\text{Level 1 (personal): } \text{logit}(p_{ij}) = \beta_{0j} + \beta_1 \text{AGE19}_{ij} + \beta_2 \text{AGE36}_{ij} + \beta_3 \text{FEMALE}_{ij} + \beta_4 \text{MLTIBIRTH}_{ij} + (\beta_5 \text{PRETERM\_BIRTH}_{ij})$$

$$\text{Level 2 (neighbourhood): } \beta_{0j} = \gamma_{00} + \gamma_{01} \text{ETHNICITY} + \gamma_{01} \text{LOW\_SES} + v_{0j}. \quad (5.5)$$

Neighbourhood ethnic variables were fit into Equation (5.5) separately one at a time. The results are listed in Table 5.18.

<i>Outcome</i>	<i>Effect</i>	<i>Estimate</i>	<i>Standard Error</i>	<i>t Value</i>	<i>Pr &gt;  t </i>
<b>Preterm births</b>	NO_CITIZEN	0.04243	0.3195	0.13	0.8944
	IMMIGRANTS	0.2866	0.1607	1.78	0.0744
	VIS_MINO	0.1388	0.1530	0.91	0.3644
	ABORIGINAL	0.1198	0.6556	0.18	0.8550
	NONOFFLANG	0.1436	0.1331	1.08	0.2806
<b>LBW</b>	NO_CITIZEN	<b>0.9204</b>	<b>0.3907</b>	<b>2.36</b>	<b>0.0185</b>
	IMMIGRANTS	<b>0.5560</b>	<b>0.2015</b>	<b>2.76</b>	<b>0.0058</b>
	VIS_MINO	<b>0.6135</b>	<b>0.1896</b>	<b>3.23</b>	<b>0.0012</b>
	ABORIGINAL	0.3558	0.8235	0.43	0.6657
	NONOFFLANG	<b>0.4966</b>	<b>0.1654</b>	<b>3.00</b>	<b>0.0027</b>

Table 5.18: Associations of ethnic composition and adverse birth outcomes after controlling for neighbourhood low SES

For Aboriginal populations, the above results show that the effects of neighbourhood-level Aboriginal composition on both LBW and preterm births disappeared after controlling for neighbourhood-level SES, suggesting a completely confounded effect. The mediational analysis results in Appendix C and D also show that neighbourhood-level Aboriginal composition affects

LBW and preterm births through all four personal-level pathways, namely mothers' behaviours, SES, psycho-social factors, and health status. This is very similar to the meditational effects of neighbourhood-level SES on LBW and preterm births, which, as discussed earlier, also affect mothers' personal-level risks through all four pathways. Thus, the negative impacts of neighbourhoods with high Aboriginal populations on adverse birth outcomes are largely due to the low SES of these neighbourhoods. The Aboriginal culture or inherited predisposition does not show impacts on the dependent variables. Hence, the Null Hypothesis of Hypothesis 4 specifically for the Aboriginal population, namely that *there is no direct or indirect impact of neighbourhood Aboriginal composition on adverse birth outcomes or on related health behaviours during pregnancy*, cannot be rejected. The two alternative hypotheses of Hypothesis 4 for Aboriginal populations are therefore rejected.

For other neighbourhood-level ethnic variables, similar patterns can also be observed for preterm births in neighbourhoods with a high proportion of minority populations. All impacts of neighbourhood ethnic variables on preterm births disappeared after controlling for neighbourhood SES. However, Appendix C shows that the mediation patterns of these risks are different than those for neighbourhoods with low SES. Although neighbourhood ethnic variables are associated with low neighbourhood SES, the associations are moderate and the negative impacts of these factors on preterm births are mainly through lower maternal SES and increased maternal health risks. There are decreased adverse health behaviours for mothers living in neighbourhoods with more immigrants and non-English speaking people, and a reduced association of mother's poor health with preterm births. Since non-English speaking mothers have a lower likelihood of preterm births (Table 5.9), the positive impacts may be due to better health behaviours and some degree of within-group support, because the strong within-group social ties of the minority groups within a neighbourhood may help to regulate a mother's health behaviours and increase mutual support. This phenomenon does show some positive impacts of group cultural influences to pregnancy outcomes.

Despite these positive factors, there is also increased psycho-social stress for mothers living in neighbourhoods with a high percentage of non-Canadian citizens. This can be explained by the associated lower social positions of residents. It may also be explained by a lack of between group integration and lack of social capital for the neighbourhood as a whole. The local within-group social ties or social networks do not necessarily translate into high social capital at the neighbourhood level. Disadvantaged urban neighbourhoods with strong within-group social ties may actually impede social integration (Wilson, 1996). This high degree of within-group social integration isolates the disadvantaged groups from contacts in broader mainstream society and makes them have low levels of informal social control over their immediate environment. The lack of between group interactions

between minority and majority ethnic groups may reduce mutual understanding and support, information sharing, and social integration at the neighbourhood level. Hence, the most disadvantaged groups in these neighbourhoods may not necessarily be ethnic minorities in general, but rather the low income ethnic majorities due to their lack of both within-group social ties and between-group social integration. Thus, while in general people living in such neighbourhoods have low SES and consequently have more chances to have preterm births, it may be that the non-ethnic group residents suffer more from increased psycho-social stress, poorer health, and adverse health behaviours, all of which lead to the increased chance of having preterm births.

For LBW, the patterns of neighbourhood ethnic impacts are somewhat different. Table 5.18 shows that, after controlling for neighbourhood-level SES, neighbourhood ethnic minority composition still shows positive impacts on adverse birth outcomes, although neighbourhood-level SES does show some mediational effects. While the mediational analysis results in Appendix D show that neighbourhoods with a high percentage of non-Canadian citizens and non-English speaking population have similar pathway patterns with neighbourhood low SES in terms of their impacts on LBW, neighbourhoods with a high percent of immigrants and visible minorities do behave differently. Despite the fact that neighbourhood immigrant and visible minority variables are associated with low neighbourhood-level SES, the behavioural pathway that is presented for low neighbourhood-level SES disappears for these two ethnic variables. This may imply some positive cultural influence among immigrants and visible minorities. The other phenomenon is that although neighbourhoods with a high percentage of non-English speaking residents are positively associated with LBW, the association does not hold at the personal level. As explained earlier for preterm births, this neighbourhood risk may actually be due to non-ethnic group residents who suffer from both low SES and the lack of inter-group social interactions. In addition, after controlling for neighbourhood-level SES and personal-level risk factors, neighbourhood minority ethnic composition is still positively associated with LBW. While the possibility that this association is due to some inherited predispositions to LBW cannot be eliminated since genetic impacts at the personal level cannot be controlled, it is also highly possible that the direct impacts may be due to some unidentified environmental risks associated with ethnic composition that directly cause adverse birth outcomes.

Although the lack of personal-level genetic information and physical environmental risks have constrained the ability for further analysis, some conclusions on the impacts of neighbourhood-level ethnic composition (excluding north American Aboriginal populations) still can be suggested based on the results discussed above. Given the existence of direct neighbourhood ethnic impacts on LBW after controlling for neighbourhood SES (Table 5.18) and the indirect cultural impacts on both LBW and preterm births discussed above, the Null Hypothesis,  $H4_0$ , namely that *there is no direct or*

indirect impact of neighbourhood ethnic composition on adverse birth outcomes or on related health behaviours during pregnancy, is not supported.

The first alternative hypothesis, *H4<sub>1</sub>*, namely that neighbourhood ethnic composition affects mothers' adverse birth outcomes by changing mothers' health behaviour during pregnancy through cultural influences within their neighbourhoods, cannot be rejected due to potential positive impacts of group cultural influences discussed above. However, since some neighbourhood ethnic variables, such as non-Canadian citizen and non-English speaking population, did not show potential cultural impacts on LBW, this alternative hypothesis may only be moderately supported for LBW births.

The second alternative of Hypothesis 4 involves the tests of both direct ethnic impacts on adverse birth outcomes and indirect ethnic impacts on adverse health behaviours during pregnancy. For the direct ethnic impacts, if an inherited predisposition of specific group members to adverse birth outcomes exists, the association of ethnic composition with adverse birth outcomes at the neighbourhood-level should still exist after controlling for confounders and personal-level ethnic status should be positively associated with adverse birth outcomes. However, the neighbourhood-level association disappeared for preterm births after controlling for neighbourhood-level SES and the personal-level ethnic variable (non-English speaking mother) is not associated with LBW.

For indirect ethnic impacts, if an inherited predisposition of specific group members to adverse health behaviours during pregnancy exists, personal-level ethnic status should be positively correlated with maternal adverse health behaviours or there should be a modification effect of maternal ethnic status on health behaviours. The earlier test on personal-level non-English speaking mothers showed that both of these assumptions were not supported. Thus, the second alternative hypothesis *H4<sub>2</sub>*, namely that neighbourhood ethnic composition affects mothers' adverse birth outcomes due to the inherited predisposition of specific group members to adverse birth outcomes or to adverse health behaviours during pregnancy, is not supported. However, given the lack of information on genetic compositions and personal-level racial/ethnic status (only the non-English speaking mother variable was used), potential direct or indirect genetic impacts on adverse birth outcomes that are associated with specific ethnic groups cannot be effectively analyzed. Further research is needed to make more solid conclusions.

### **5.3.5 Hypothesis 5: Global vs. Local Influences**

Within the theoretical frameworks discussed in Chapter 3, the above analyses have described the overall impacts of neighbourhood-level risks on adverse birth outcomes and suggest significant impacts of different neighbourhood characteristics on adverse birth outcomes. However, the analyses are global in terms of identifying universally applicable associations. Due to local clustering effects

or the complexity of local contexts, some of the exposures may only be harmful to certain neighbourhoods and some spatial variations may not be appropriately modeled. Local “hot spots” of adverse birth outcomes may exist and personal and neighbourhood-level risks may affect adverse birth outcomes differently from neighbourhood to neighbourhood. It is therefore necessary to examine local variations in terms of both outcomes and effects so that health interventions may be conducted more purposefully and objectively. These influences are now examined.

#### 5.3.5.i Local spatial clustering analysis

Due to different levels of social and environmental impacts, clustering of adverse birth outcomes may occur at several spatial scales, from the closest living environment (neighbourhoods) through municipal, regional, provincial, and the national spatial levels. Given the advantage of identifying clusters at different spatial scales, the Bayesian spatial scan statistic described in Equation (4.26) was used to analyze local spatial clustering. Since the Bernoulli model and the Poisson model closely approximate each other when events have a small probability of occurrence, it does not make any difference which of these models is used given the small occurrence rates of both LBW and preterm births in the study area. The Poisson model was therefore used at the CDA level and fitted by the SaTScan software. For the analysis of LBW births, the total number of observations in each CDA was taken to be the number of singular full-term births at the CDA and the observed occurrence of LBW births was the number of singular full-term LBW births in the CDA. For the analysis of preterm births, the total number of observations in each CDA was calculated as the number of singular births in the CDA and the observed occurrence of preterm births is the number of singular preterm births in the CDA. Since the point of interest was to identify high risk clusters of adverse birth outcomes, only high risk clusters were scanned and statistically tested.

The resulting maps for the clustering of high risks for preterm births LBW births are presented in Figures 5.11 and 5.12 respectively. A Monte Carlo simulation was used after the values of the test statistic were calculated. Altogether 9999 replications of the data set were generated under the null hypothesis that the probabilities of adverse birth outcomes within and outside the test zone are the same. The distribution of the simulated replications shows that the identified clusters are all significant at the 95% confidence level.

The values shown in Figures 5.11 and 5.12 identify the relative risks of the incidence rates within the buffer areas compared to the incidence rates outside the buffer areas. All clusters were found to exist within the health unit of Windsor and Essex County. The biggest cluster was the whole health unit of Windsor and Essex County, which is at risk for both preterm births (relative risk 1.2)

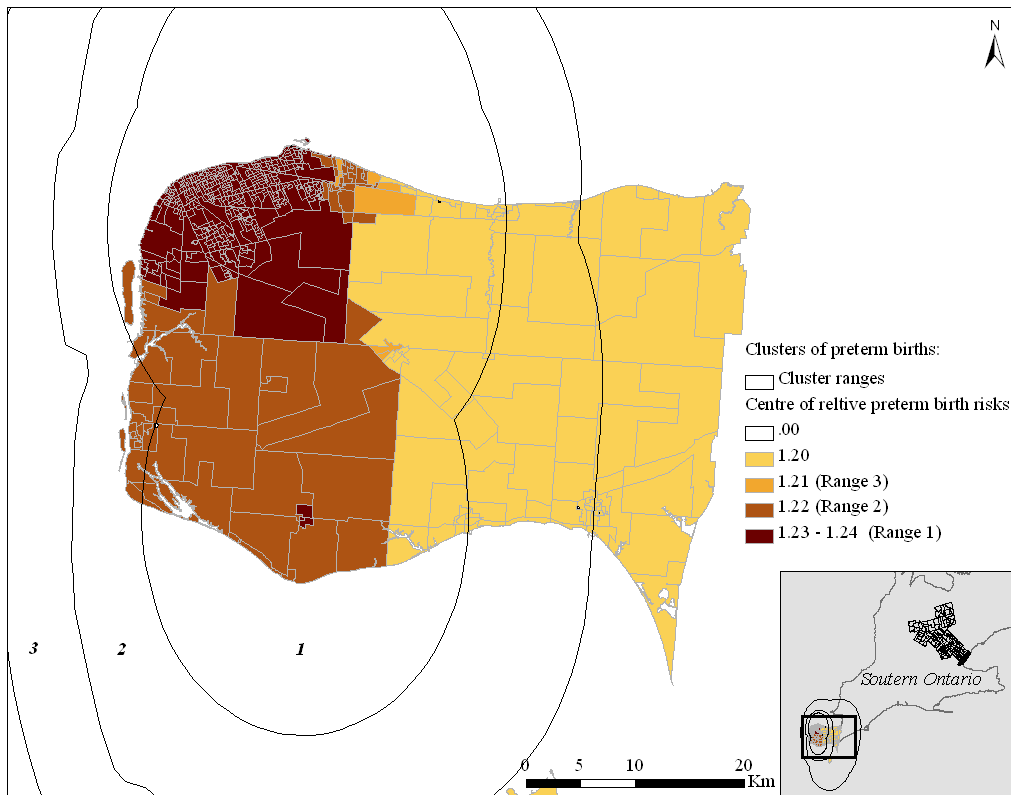


Figure 5.11: Preterm birth clusters in the study region

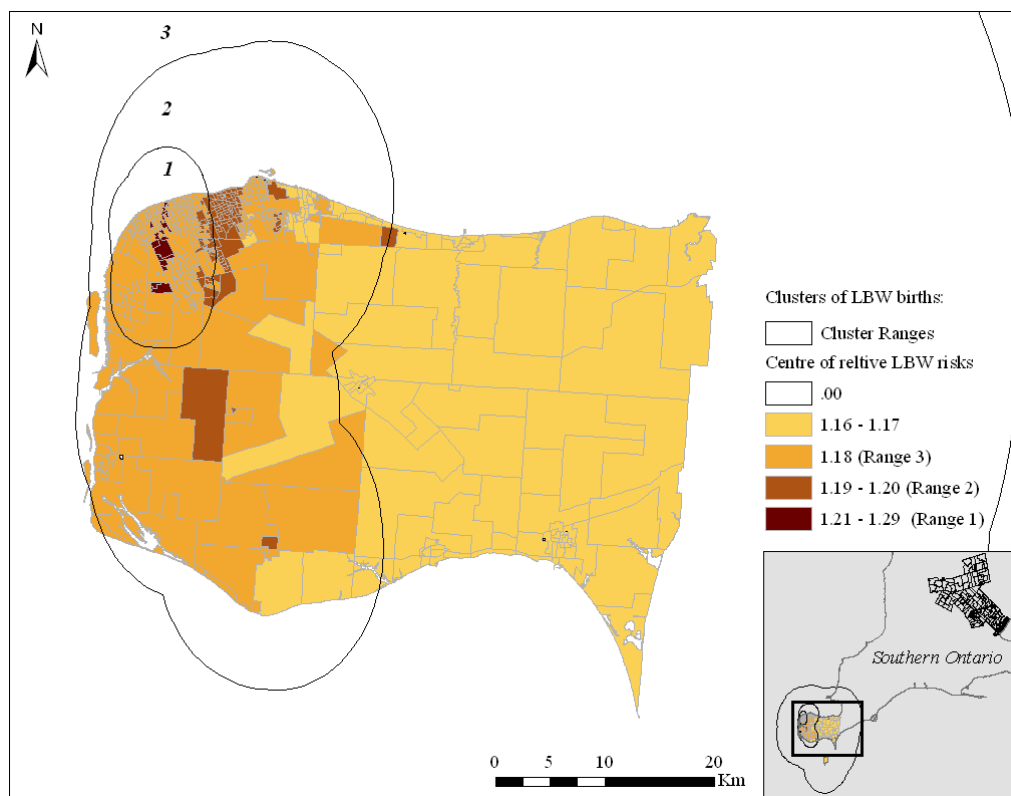


Figure 5.12: LBW birth clusters in the study region



and LBW (relative risk 1.16-1.17) compared to the other two health regions. This is not surprising based on the descriptive statistics presented in Tables 5.3 and 5.4. The spatial scan statistic proves that the higher LBW and preterm birth rates in this health region are statistically significant. The clusters with the highest risks for both LBW and preterm births were centred in the City of Windsor, with a relative risk of 1.2-1.29 for LBW, and a relative risk of 1.23-1.24 for preterm births. The range buffers in both maps represent affected areas of these highest risks. The range 2 buffers indicate clusters with moderate relative risks for both preterm births and LBW centred at the western half of the region.

Thus, although the affected ranges of the clusters are clearly different, preterm birth and LBW clusters tend to have similar patterns, suggesting common environmental impacts on the clusters of both measures of adverse birth outcomes. The concentric zone patterns of both clusters also suggest that there may exist a potential source of environmental risk close to the City of Windsor, and a distance decay effect of this potential risk source is observed as shown in Figures 5.11 and 5.12. Further investigation is required to seek the exact causes of these clusters. This finding proves the existence of local “hot spots” which may not be addressed by a global intervention on personal risks. Hence, specific area-based procedures need be used to address this issue.

#### 5.3.5.ii Disease mapping

Given the existence of spatial clustering of adverse birth outcomes, it is desirable to depict the spatial distribution of these outcomes to identify which neighbourhoods have higher risks relative to others. This is important both from a policy and planning perspective in terms of intervention to remediate the causes of the adverse birth outcome relationships that the mediational analysis has revealed.

Spatially structured random effect models have shown advantages in controlling for confounding factors at different influence levels, including personal and neighbourhood levels, and mapping the spatial distribution of the residual effects. A convolution prior model similar to Equation (4.24) was therefore used to map the spatial distribution of LBW and preterm births. The model is constructed as:

*Adverse birth outcomes*  $\sim$  *binary* ( $p_{ij}$ )

$$\text{Level 1 (personal):} \quad \text{logit}(p_{ij}) = \beta_{0j} + \beta_1 \text{AGE19}_{ij} + \beta_2 \text{AGE36}_{ij} + \beta_3 \text{FEMALE}_{ij} + \beta_4 \text{MLTIBIRTH}_{ij} \\ + (\beta_5 \text{PRETERM\_BIRTH}_{ij})$$

$$\text{Level 2 (neighbourhood):} \quad \beta_{0j} = \gamma_{00} + u_{0j} + v_{0j}. \quad (5.6)$$

LBW and preterm births were fitted into this model separately and statistics calculated using WinBUGS 1.4.3 software. Given its simplicity, the intrinsic conditional autocorrelation (CAR) model discussed in Chapter 4.3.1 was used as the prior distribution of  $u_{0j}$  to describe the spatial

autocorrelation of the structured neighbourhood-level random effect. However, since the spatial autocorrelation matrix, with elements  $c_{jk}$ , is given based on certain assumptions of spatial dependency rather than calculated by the model, the selection of neighbourhood structures may have potential impacts on the calculation of the neighbourhood-level random effects.

The CAR model based on a contiguity neighbourhood structure assumes that  $u_{0j} \sim \text{Normal}(\bar{u}_{0j}, v_{u0j})$  where  $\bar{u}_{0j} = \frac{\sum_{k \text{ in } N_j} u_{0k}}{n_j}$ , and  $v_{u0j} = \frac{\text{variance}(u_0)}{n_j}$ . This is to say that the expected spatially structured random effect at areal unit  $j$  is the average of values of areal units that are adjacent to  $j$ , and the variance is inversely related to the number of unit  $j$ 's adjacent areal units. Every adjacent areal unit of  $j$  has the same correlation with  $j$  and is therefore equally weighted for its contribution to the mean of  $j$ . This assumption holds true when only immediate neighbourhood environments have impacts on adverse birth outcomes and interrelations are present only between adjacent neighbourhoods.

For the small-area analysis of adverse birth outcomes, the above assumption may or may not hold true depending on the real spatial impacts. For instance, spatial influences may occur not only through adjacent neighbourhoods, but also through a function of distance. As noted in Chapter 4, neighbourhoods in the study area were represented by census dissemination areas, which are not the same size. CDAs are small within cities and large in rural areas. Hence, the distances between the centroids of two adjacent neighbourhoods can vary from several hundred metres to over 10 kilometres. Neighbourhood structures defined by contiguous areal units may not be appropriate if the spatial relation is a function of distance. The CAR model in the distance function setting is equivalent to  $u_{0j} \sim \text{Normal}(\bar{u}_{0j}, v_{u0j})$  where  $\bar{u}_{0j} = \frac{\sum_{j \neq k} w_{kj} u_{0k}}{\sum_{j \neq k} w_{kj}}$ , and  $v_{u0j} = \frac{\text{variance}(u_0)}{\sum_{j \neq k} w_{kj}}$ . The weight  $w_{kj}$  is defined as:

$$w_{kj} = \begin{cases} \frac{1}{d_{kj}^\alpha} & \text{if } k \text{ and } j \text{ are within a certain distance, } \delta \\ 0 & \text{otherwise,} \end{cases}$$

where  $d_{kj}$  is the distance between unit  $k$  and  $j$ , and  $\alpha$  determines how rapidly the correlations diminish as distance increases. The value of  $\alpha$  is often set to 1 or 2.

To compare which neighbourhood structure better describes the spatial autocorrelations of adverse birth outcomes, Equation (5.6) was fitted separately for preterm births using WinBUGS based on the above two neighbourhood structures. The WinBUGS code can be found in Appendix E.

Appendix F shows the summary statistics and plots for convergence diagnostics of the model for preterm birth analysis. Convergence was first monitored by visual examination of the history plots of

the sample parameter values versus iteration. The distribution quickly stabilized after several iterations. It can be observed that, after several runs, sample values are randomly distributed above and below a stable mean value for each parameter and the two chains from different initial values appear to be overlapping with each other. The autocorrelation plots show independence of samples after a small number of lags. The Brooks, Gelman and Rubin statistics also show that both the within- and between-chain variances stabilized over the second half of the two chains to the same value and the ratio  $R$  (red lines) was stabilized to 1. Therefore, it can be stated with confidence that convergence has been achieved for this model.

The choice of using the hyperprior distributions of the parameters of the fixed effects is based on a preliminary fit of equation (5.1) calculated with SAS. The results showed that all values of the fixed effects were within the range of -3 to 3. Thus the hyperprior distributions were set to have normal distributions with means of 0 and variances of 100. This setting allows enough range to cover the true parameter values and at the same time narrows down the parameter prior distributions to have a relatively small variation so that convergence can be quickly achieved. For sensitivity testing, two prior distributions, namely Gamma (0.01, 0.01) and Gamma (0.5, 0.0005), were selected as possible hyperpriors of the variances of the random effects. The later performed better with a smaller DIC value and was chosen to be the model for generating posterior distributions for inferences.

After convergence was achieved, 10000 dependent sample values from each chain were generated from the posterior distribution to approximate the distribution. The Monte Carlo error for each parameter was less than 5% of the sample standard deviation. Point estimates were then obtained from the simulated sample values.

For the inverse distance neighbourhood structure,  $\alpha$  was set to 1 so that a gradually reduced autocorrelation may be captured rather than a high dependence on the immediate neighbourhoods, and  $\delta$  was set to 10500 metres, which is the largest centroid distance between two neighbouring units, so that each areal unit would have at least one neighbour. The model parameter estimates are listed in Tables 5.19 and 5.20 respectively.

While the fixed parameter estimates were similar for these two prior settings of neighbourhood structures, the random effects were somewhat different. Specifically, the Deviance Information Criterion (DIC) (Spiegelhalter et al., 2002) was used to assess model complexity and compare the two models. DIC is a Bayesian method that WinBUGS calculates for model comparison. It is measured by the goodness of fit for the data plus the model complexity. The model with the smallest DIC is the best model predicting a dataset with same structure as the currently observed data. It can be observed from the above results that the DIC value for the model with a contiguous neighbourhood structure is somewhat smaller, although the difference is quite small. It can also be observed that the total

variances for the random effects are larger for the model with the inverse distance neighbourhood structure, indicating a lower goodness of fit. It also shows a positive correlation between  $u_0$  and  $v_0$  for the model with the inverse distance neighbourhood structure since the variance of  $u_0+v_0$  is greater than the sum of variance  $u_0$  and variance  $v_0$ . This indicates that  $u_0$  and  $v_0$  may not be appropriately separated by this model.

<i>Parameter</i>	<i>Mean</i>	<i>Standard Deviation</i>	<i>2.50%</i>	<i>Median</i>	<i>97.50%</i>
$\gamma_{00}$	-2.780	0.024	-2.828	-2.780	-2.733
$\beta_1$	0.233	0.065	0.105	0.234	0.358
$\beta_2$	0.229	0.043	0.145	0.230	0.312
$\beta_3$	-0.132	0.028	-0.188	-0.131	-0.077
$\beta_4$	2.967	0.043	2.883	2.967	3.051
Variance $u_0$	0.020				
Variance $v_0$	0.005				
Variance $u_0+v_0$	0.027				
Total DIC	39131.8				

Table 5.19: Parameter estimates for preterm births based on a contiguity neighbourhood structure

<i>Parameter</i>	<i>Mean</i>	<i>Standard Deviation</i>	<i>2.50%</i>	<i>Median</i>	<i>97.50%</i>
$\gamma_{00}$	-2.794	0.028	-2.851	-2.793	-2.742
$\beta_1$	0.245	0.065	0.117	0.245	0.371
$\beta_2$	0.225	0.043	0.140	0.224	0.310
$\beta_3$	-0.133	0.028	-0.188	-0.133	-0.077
$\beta_4$	2.976	0.043	2.892	2.975	3.061
Variance $u_0$	0.0357				
Variance $v_0$	0.0036				
Variance $u_0+v_0$	0.0439				
Total DIC	39150.8				

Table 5.20: Parameter estimates for preterm births based on an IDW neighbourhood structure

In addition, if the spatial structure is perfectly captured by  $u_0$ , the unstructured random effect  $v_0$  should be spatially uncorrelated. To examine this, the global Moran's  $I$  statistics are given next to test the spatial dependency of  $v_0$  values that result from both models. The results show that  $v_0$  values calculated by the model using the contiguity neighbourhood structure do not have spatial autocorrelation ( $I = -0.004$ ,  $p\text{-value} = 0.42$ ), but  $v_0$  values calculated by the model using the inverse distance neighbourhood structure do show statistically significant spatial dependencies ( $I = 0.008$ ,  $p\text{-value} = 0.027$ ). This result shows that the contiguity neighbourhood structure performs better than the

inverse distance neighbourhood structure for measuring spatial dependencies of adverse birth outcomes.

The distributions of  $u_0$  and the total random effects,  $u_0+v_0$  for these two models are mapped in Figure 5.13 through Figure 5.16. It can be visually observed from these figures that the contiguity neighbourhood structure captures the spatial dependence nicely (Figure 5.13), and the inverse distance neighbourhood structure also captures some noise (Figure 5.15), which should be modeled by  $v_0$ . Bayesian testing is also provided for  $u_0$  at the 5 percent significance level for each CDA, based on the posterior distribution of  $u_0$ . For a given CDA, if over 95 percent of the simulated samples within it based on the posterior distribution of  $u_0$  are positive, the CDA has a significant positive  $u_0$  value. On the other hand, if over 95 percent of the simulated samples are negative, the CDA has a significant negative  $u_0$  value. The CDAs outlined with blue lines in Figures 5.13 and 5.15 have significant positive or negative  $u_0$  values.

It can be observed from both Figures 5.13 and 5.16 that high risks of preterm births exist within the city of Windsor. This is consistent with the earlier identified highest risks by the scan statistic. However, since this random effect model can only identify risks at the CDA level, larger scale clusters identified by the scan statistic cannot be captured. In Figure 5.13, the contiguity neighbourhood structure also captures two significant low risk areas. One is in the centre of Wellington County and the other is at the east corner of Halton Region. To be able to capture the overall variation of preterm births, the total random effects  $u_0+v_0$  were mapped. Comparing Figure 5.14 with Figure 5.16, while the overall patterns are similar, there are some local differences, indicating the selection of neighbourhood structures does have some non-ignorable influences on the estimation of spatial variations of preterm births.

Given the above discussed advantage, the contiguity neighbourhood structure was selected for this and later analysis and the corresponding results were used for interpretation. Based on Figure 5.15, the total random effect values in the map combined with the intercept (-2.8) represent the log relative risks,  $\text{Log}(p/I-p)$ , of preterm births.

The total area level effects range from -0.4(-2.8) to 0.493(-2.8). This means that after controlling for mothers' age, babies' sex, and multiple births, the probabilities of having preterm births range from 3.9% to 9%, with an overall mean of 5.7%. The dark brown areas located within the city of Windsor depict a cluster of high preterm birth rates with probabilities ranging from 6.7% to 9%. Compared to the average probability after controlling for other factors (5.7%), women living in these areas have 1.2 to 1.6 times the risk of having a preterm birth baby than the average risk of women living in the study area.

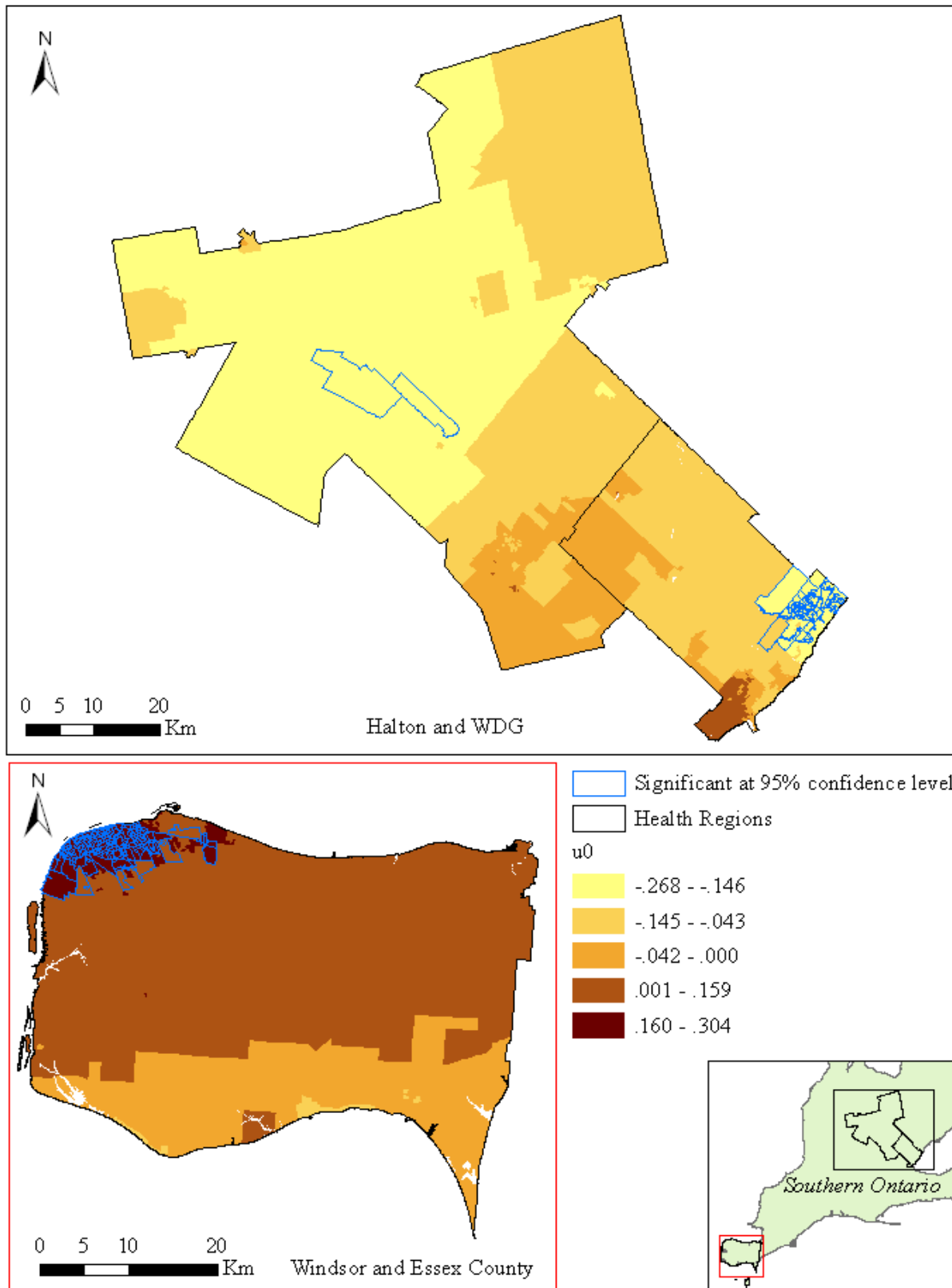


Figure 5.13: Distribution of spatially correlated random effects  $u_0$  for preterm births based on contiguity neighbourhood structure

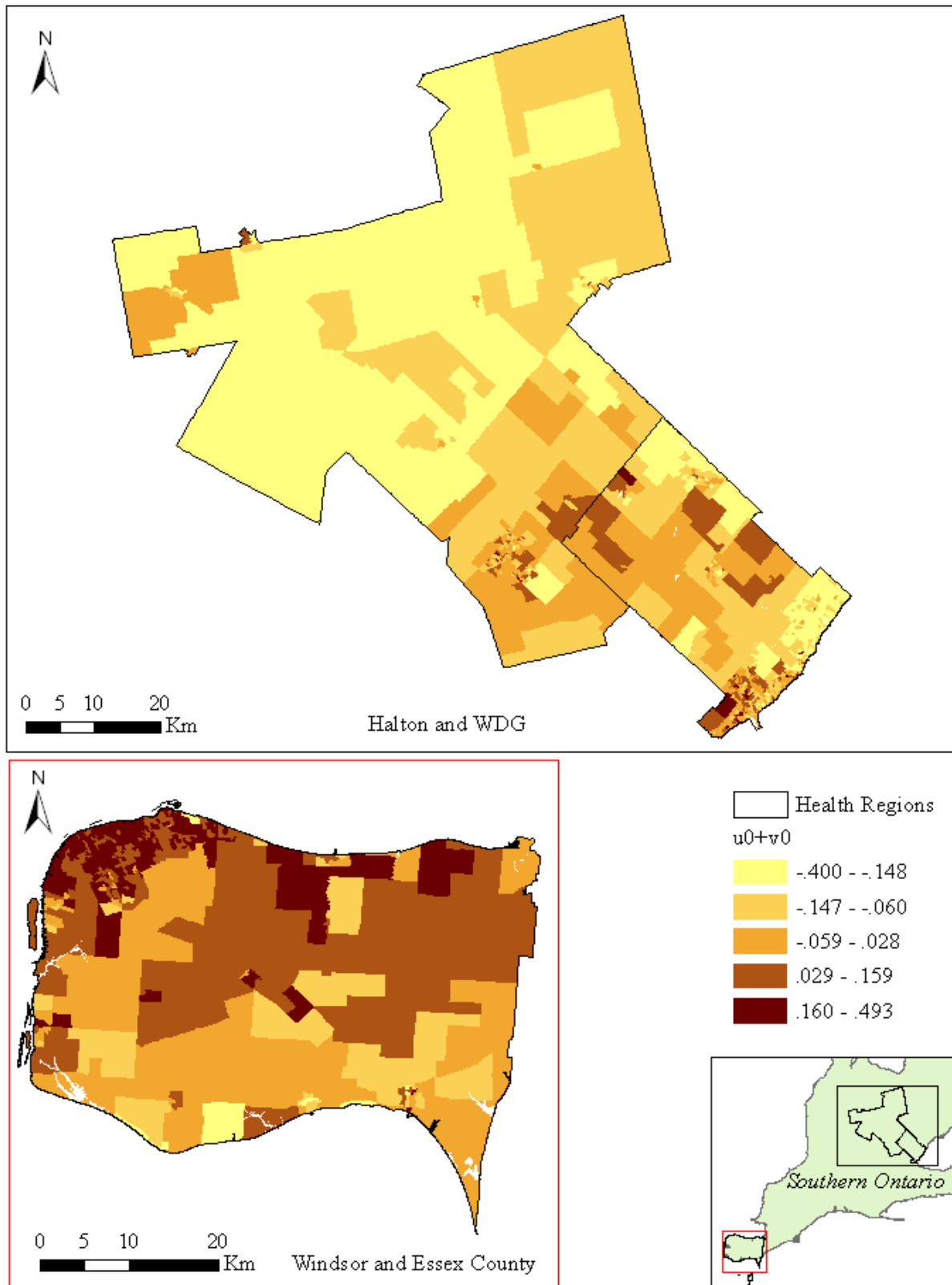


Figure 5.14: Distribution of total random effects  $u_0+v_0$  for preterm births based on a contiguity neighbourhood structure

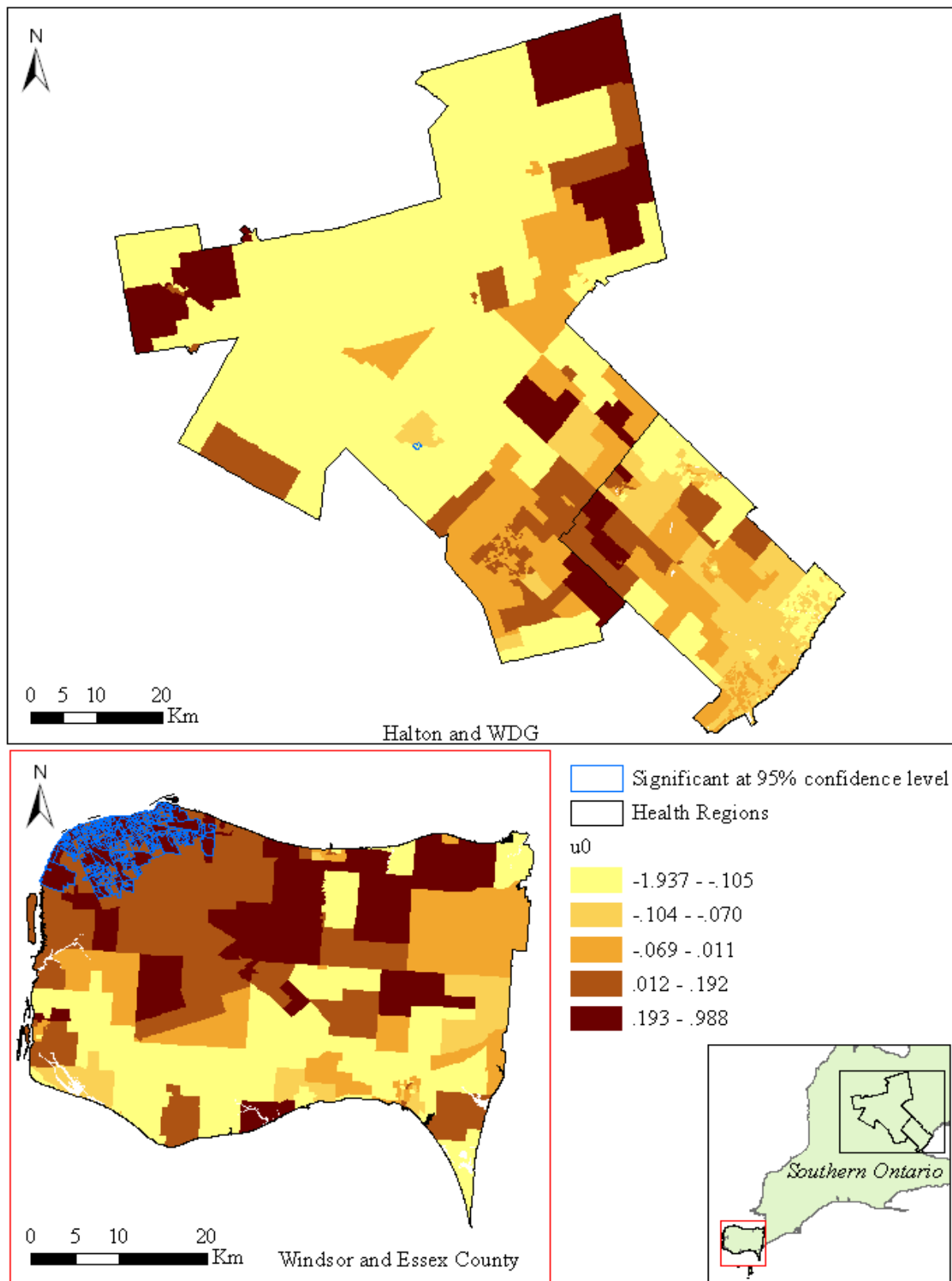


Figure 5.15: Distribution of spatially correlated random effects  $u_0$  for preterm births based on an inverse distance neighbourhood structure



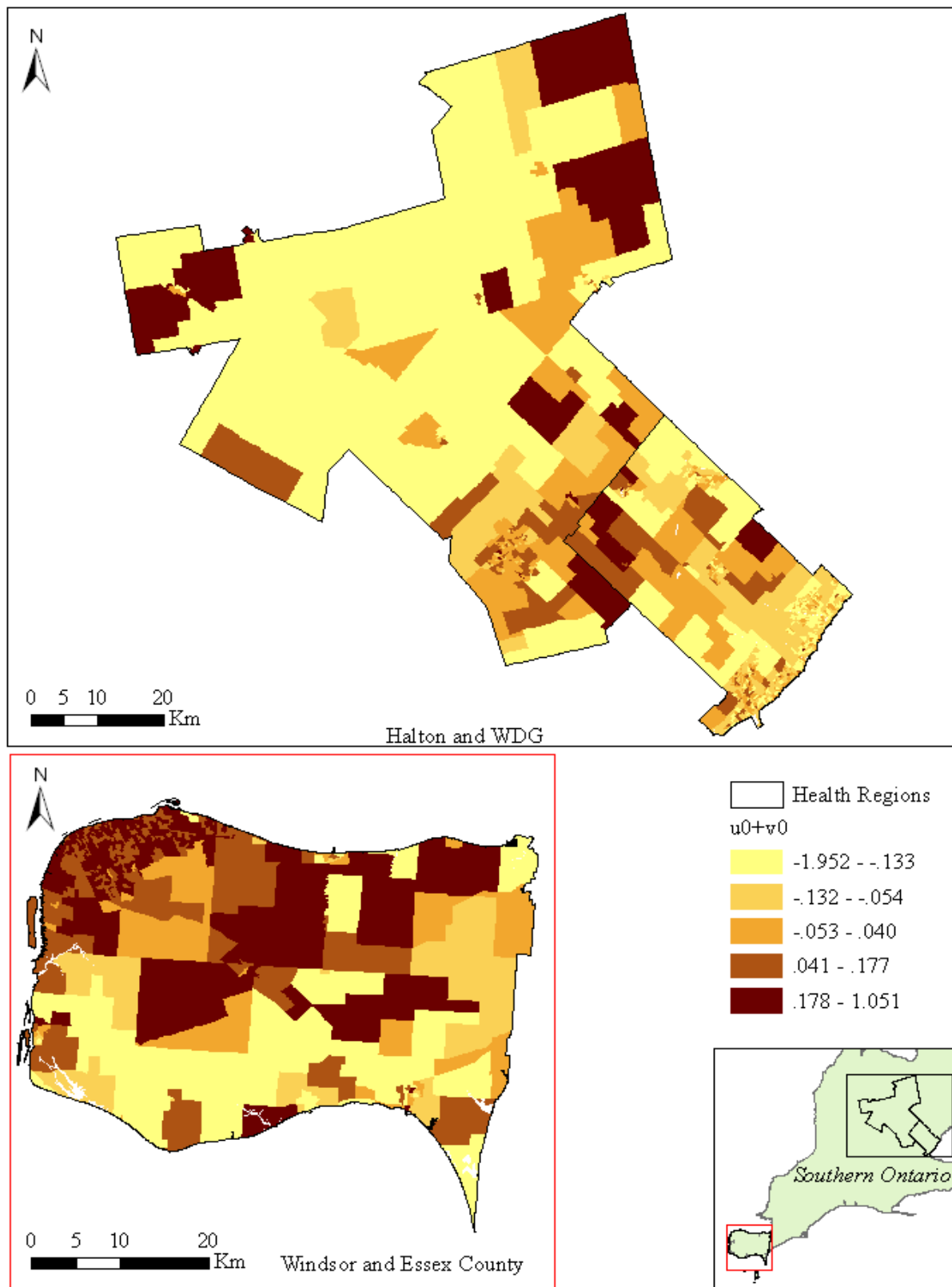


Figure 5.16: Distribution of total effects  $u_0 + v_0$  for preterm births based on an inverse distance neighbourhood structure

The random effect models were then similarly constructed and calculated for LBW births. The results also show that the model using the contiguity neighbourhood structure performs better and this model was therefore used for interpretation. The parameter estimates based on a contiguity neighbourhood structure are listed in Table 5.21.

<i>Parameter</i>	<i>Mean</i>	<i>Standard Deviation</i>	<i>2.50%</i>	<i>Median</i>	<i>97.50%</i>
$\gamma_{00}$	-4.117	0.03769	-4.191	-4.117	-4.043
$\beta_1$	0.3921	0.07834	0.2364	0.3925	0.5426
$\beta_2$	0.1853	0.05603	0.07446	0.1849	0.2961
$\beta_3$	0.2355	0.03652	0.1639	0.2358	0.3069
$\beta_4$	2.066	0.05815	1.952	2.066	2.18
$\beta_5$	3.612	0.03834	3.537	3.612	3.687
Variance $u_0$	0.016				
Variance $v_0$	0.007				
Variance $u_0+v_0$	0.028				

Table 5.21: Parameter estimates for preterm births based on a contiguity neighbourhood structure

The distributions of  $u_0$  and  $u_0+v_0$  are shown in Figures 5.17 and 5.18 respectively. Figure 5.17 shows that significant high risks of LBW births are still found within the city of Windsor. The average LBW rate after controlling for mothers' age, babies' sex, multiple births, and preterm births is 1.64%, and varies from 0.97% to 3.2%. The dark brown areas show clusters of high LBW rates, representing 1.2 to 2 times the risk of giving birth to under-weight babies. In general, urban areas also have relatively higher LBW rates than rural areas. Tables 5.19 and 5.21 also show that about 80% and 70% of the total random effects of preterm births and LBW births are spatially dependent, indicating non-ignorable local spatial impacts.

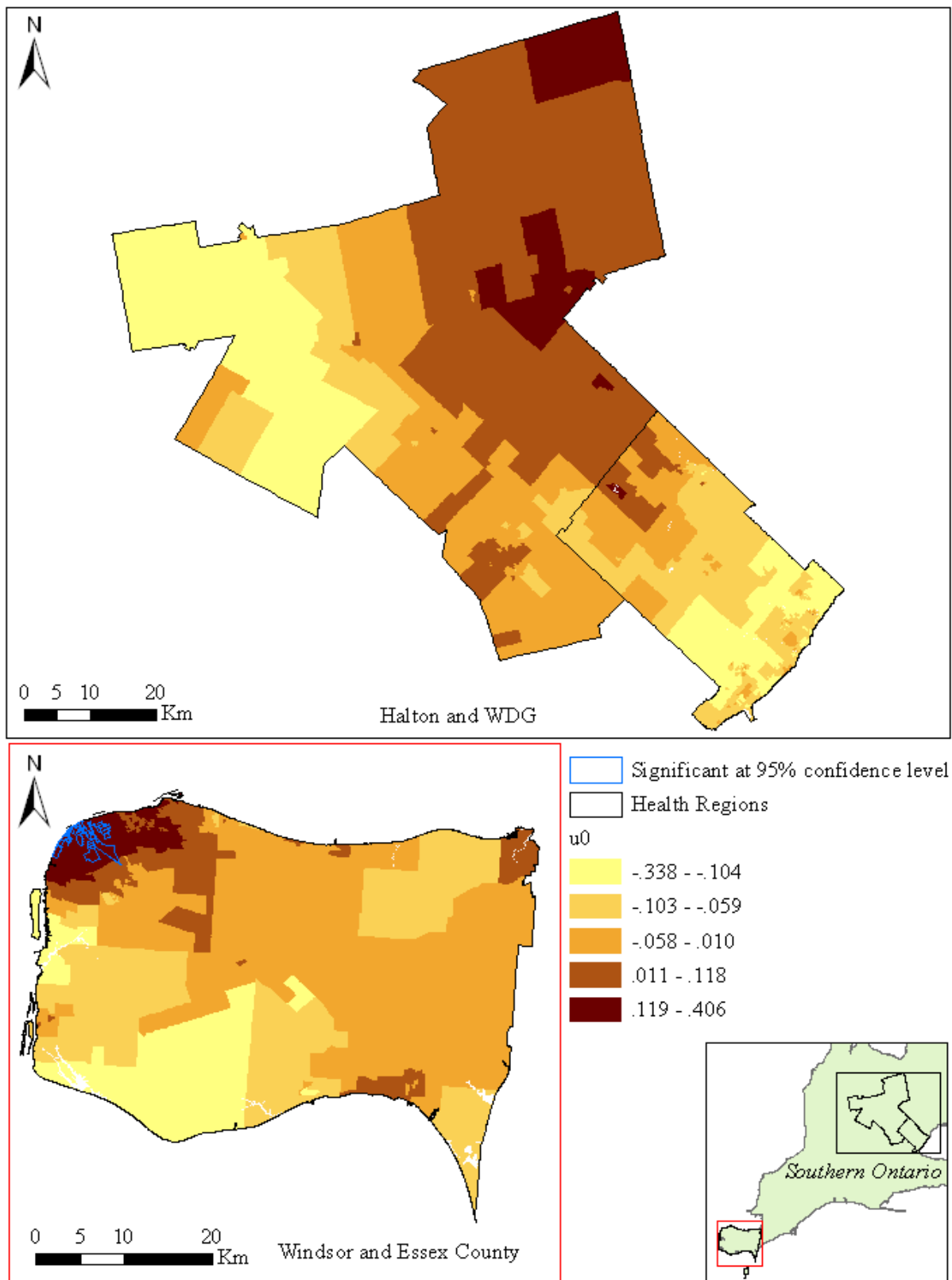


Figure 5.17: Distribution of spatially correlated random effects  $u_0$  for LBW births based on an contiguity neighbourhood structure

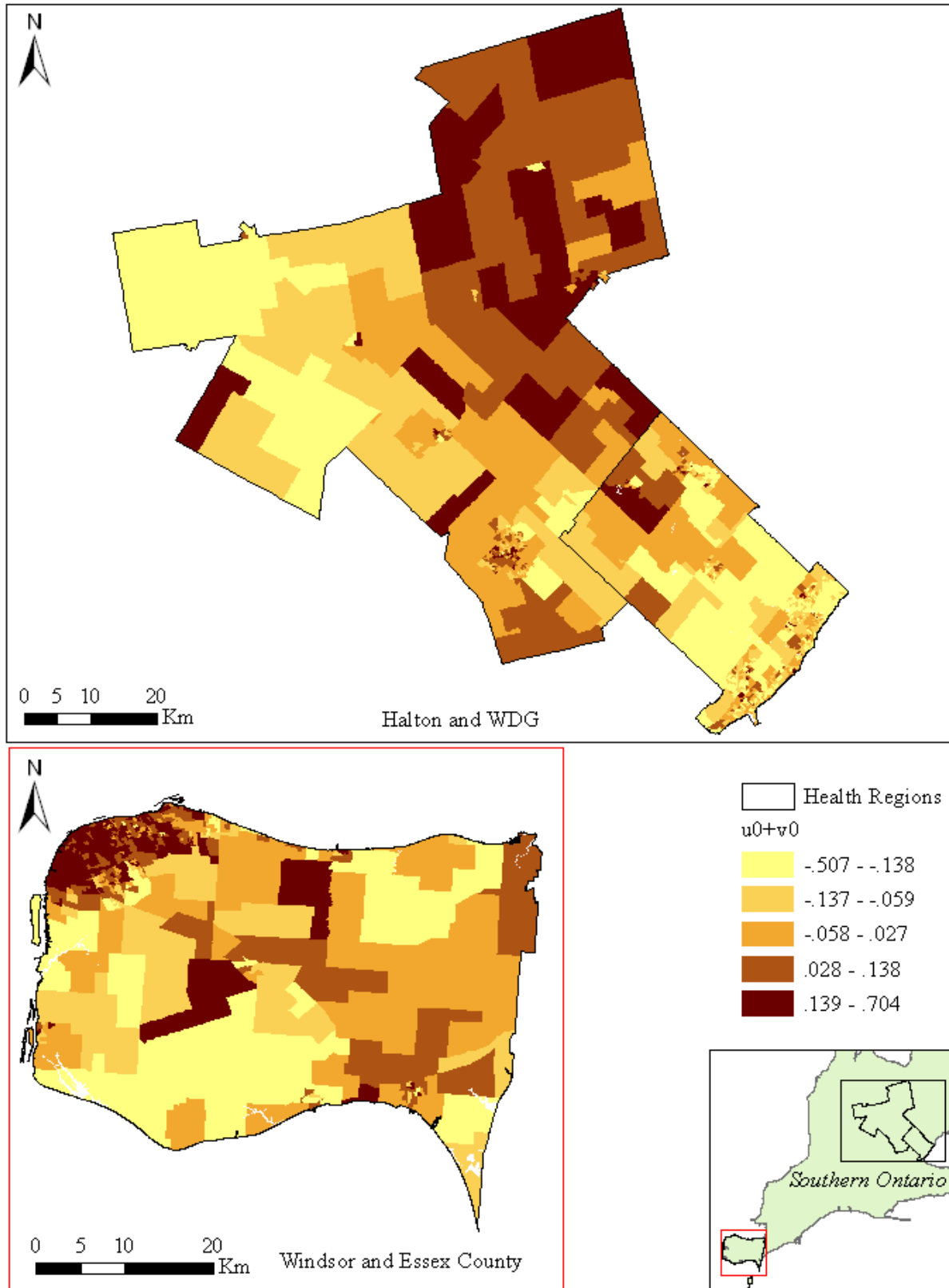


Figure 5.18: Distribution of total random effects  $u_0 + v_0$  for LBW births based on a contiguity neighbourhood structure

### 5.3.5.iii Spatial variation of personal risk factors

In addition to birth outcomes, the impacts of risk factors on birth outcomes may also differ from place to place due to local influences. To test whether the impacts of personal risk factors on adverse birth outcomes are universal, or are spatially varied, a Bayesian model was constructed similar to Equation (4.28) as:

*Adverse birth outcomes*  $\sim$  *binary* ( $p_{ij}$ )

Level 1 (personal):  $\text{logit}(p_{ij}) = \beta_{0j} + \beta_1 \text{AGE19}_{ij} + \beta_2 \text{AGE36}_{ij} + \beta_3 \text{FEMALE}_{ij} + \beta_4 \text{MLTIBIRTH}_{ij} + (\beta_5 \text{PRETERM\_BIRTH}_{ij}) + \alpha_{1j} \text{factor1}_{ij} + \alpha_{2j} \text{factor2}_{ij} + \alpha_{3j} \text{factor3}_{ij} + \alpha_{4j} \text{factor4}_{ij}$

Level 2 (neighbourhood):  $\beta_{0j} = \gamma_{00} + u_{0j} + v_{0j}$

$$\alpha_{1j} = \beta_6 + u_{1j} + v_{1j}$$

$$\alpha_{2j} = \beta_7 + u_{2j} + v_{2j}$$

$$\alpha_{3j} = \beta_8 + u_{3j} + v_{3j}$$

$$\alpha_{4j} = \beta_9 + u_{4j} + v_{4j}. \quad (5.7)$$

This model was used instead of using other spatial regression models (such as the Geographically Weighted Regression method) due to its flexibility to model complex multilevel impacts. It is able to incorporate risk factors at different spatial levels and model spatial impacts for both outcomes and identified risk factors. Instead of estimating the impacts of personal risk factors as having fixed effects, the model assumes the parameters of personal factors vary spatially across neighbourhoods. Neighbourhood impacts on personal risk factors can be identified through modelling spatial variations of personal risk parameters as both spatially structured and non-structured random effects. The model also uses a CAR prior and the contiguity neighbourhood structure to describe spatial dependencies. The estimates of fixed impacts and the variances of the random effects for preterm births and LBW births are listed in Tables 5.22 and 5.23 respectively. The spatially correlated random effects ( $u_0$ ,  $u_1$ ,  $u_2$ ,  $u_3$ , and  $u_4$ ), and the total CDA-level random effects ( $u_0+v_0$ ,  $u_1+v_1$ ,  $u_2+v_2$ ,  $u_3+v_3$ , and  $u_4+v_4$ ) were mapped for preterm and LBW births respectively. The resulting maps are shown in Appendix G.

Figure G.1 and Figure G.11 show clear patterns of spatial autocorrelation after controlling for personal risk factors. The areas outlined with blue lines are statistically significant high-risk areas measured at the 5% significance level according to the Bayesian posterior distribution. The clusters are still around the city of Windsor, indicating that the higher risks in these areas are not due to compositional effects of individuals, namely not due to the high concentration of low SES, poor health, psycho-socially stressed, and adverse health behaviour mothers. Tables 5.22 and 5.23 show that the spatially correlated random effects of preterm and LBW births,  $u_0$ , contribute to 46% and 19%

<i>Parameter</i>	<i>Mean</i>	<i>Standard Deviation</i>	<i>2.50%</i>	<i>Median</i>	<i>97.50%</i>
$\gamma_{00}$	-2.8290	0.0243	-2.8760	-2.8290	-2.7810
$\beta_1$	0.1910	0.0696	0.0536	0.1918	0.3258
$\beta_2$	0.2159	0.0439	0.1299	0.2160	0.3029
$\beta_3$	-0.1321	0.0287	-0.1896	-0.1317	-0.0764
$\beta_4$	2.9550	0.0434	2.8700	2.9550	3.0400
$\beta_6$	0.0675	0.0214	0.0246	0.0680	0.1081
$\beta_7$	0.0465	0.0192	0.0080	0.0469	0.0829
$\beta_8$	0.0934	0.0215	0.0501	0.0938	0.1343
$\beta_9$	0.4098	0.0200	0.3697	0.4098	0.4487
Variance $u_0$	0.0062				
Variance $v_0$	0.0056				
Variance $u_0+v_0$	0.0134				
Variance $u_1$	0.0002				
Variance $v_1$	0.0004				
Variance $u_1+v_1$	0.0007				
Variance $u_2$	0.0002				
Variance $v_2$	0.000004				
Variance $u_2+v_2$	0.0002				
Variance $u_3$	0.0008				
Variance $v_3$	0.0015				
Variance $u_3+v_3$	0.0025				
Variance $u_4$	0.0018				
Variance $v_4$	0.0023				
Variance $u_4+v_4$	0.0045				

Table 5.22: Fixed and random effects of personal parameter estimates for preterm birth

of the total random effects, *variance  $u_0$* /*variance  $u_0+v_0$* , respectively, indicating non-ignorable spatial clustering for the remaining spatial variations of preterm births. Although the spatial distributions are blurred in the total effect maps (Figures G.2 and G.12), they still show high risks of preterm and LBW births in the Windsor area. Further investigation is therefore needed to identify area-level contextual determinants that cause these high risks.

Figures G.3 and G.4 show that the overall fixed effect of the personal SES factor (Factor 1) to preterm births is 0.0675 and the random effect ranges from -0.148 to 0.139. This means that in some of the yellow areas, the effects may be inverse. It can also be observed in the figures that low maternal SES has a relatively higher impact on preterm births in the Wellington-Dufferin-Guelph

(WDG) and Halton health regions than in the Windsor and Essex Health Region. The spatially correlated random effect contributes 32% of the total variation.

<i>Parameter</i>	<i>Mean</i>	<i>Standard Deviation</i>	<i>2.50%</i>	<i>Median</i>	<i>97.50%</i>
$\gamma_{00}$	-4.1690	0.0390	-4.2460	-4.1690	-4.0930
$\beta_1$	0.2552	0.0836	0.0903	0.2556	0.4181
$\beta_2$	0.1733	0.0569	0.0621	0.1733	0.2854
$\beta_3$	0.2381	0.0369	0.1666	0.2380	0.3115
$\beta_4$	2.0780	0.0584	1.9630	2.0770	2.1910
$\beta_5$	3.6110	0.0403	3.5320	3.6110	3.6900
$\beta_6$	0.1351	0.0286	0.0776	0.1356	0.1895
$\beta_7$	0.0495	0.0287	-0.0082	0.0500	0.1045
$\beta_8$	0.1143	0.0197	0.0748	0.1147	0.1522
$\beta_9$	0.2121	0.0271	0.1579	0.2125	0.2638
Variance $u_0$	0.003307				
Variance $v_0$	0.012410				
Variance $u_0+v_0$	0.017629				
Variance $u_1$	0.000046				
Variance $v_1$	0.000649				
Variance $u_1+v_1$	0.000754				
Variance $u_2$	0.000427				
Variance $v_2$	0.003620				
Variance $u_2+v_2$	0.004183				
Variance $u_3$	0.000046				
Variance $v_3$	0.000001				
Variance $u_3+v_3$	0.000049				
Variance $u_4$	0.000612				
Variance $v_4$	0.002241				
Variance $u_4+v_4$	0.003225				

Table 5.23: Fixed and random effect of personal parameter estimates for LBW birth

In Figures G.5 and G.6, the overall fixed effect of the personal psycho-social factor (Factor 2) on preterm births is 0.0465 and the random effects range from -0.035 to 0.04. Almost all of the random effects are part of the spatially correlated random effect (96%). The figures show that maternal depression or stress has a relatively higher impact on preterm births in the Windsor and Essex Health Region and the western part of the WDG Health Region than in other places. It is especially high in the city of Windsor, the north-east part of Essex County and the west part of Wellington County.

Local investigation of these areas may be needed to identify the reasons for their higher psycho-social impacts on preterm births. Local programs to increase the social integration of these neighbourhoods and reduce the depression of women living in these areas may be valuable for reducing the clustered impacts.

Figures G.7 and G.8 show higher effects of personal behavioural risks (Factor 3) on preterm births in the Windsor and Essex health region than in the other two regions. The fixed impact is 0.0934 and the random effects range from -0.199 to 0.291. The spatially correlated random effect contributes 33% of the total variations. The spatial variation indicates that there may be an elevated risk of maternal adverse health behaviours affecting preterm births in the Windsor and Essex health region compared to the other two regions.

It can also be observed in Figures G.9 and G.10 that the median fixed effect of maternal health conditions on preterm births is 0.4098 and the random effects range from -0.215 to 0.29. The contribution of the spatially correlated random effect to the total variation is 41%. The impact of maternal health on preterm births is relatively high in Windsor and Essex County. The higher risk of maternal health conditions in these areas is statistically significant at the 95% confidence level based on the Bayesian posterior distributions. This was the only one out of the four personal risk factor coefficients that showed statistically significant spatial variation.

Figures G.3 through G.10 provide information for the local focus of interventions to reduce preterm births, and confirm the impacts of space on the inequalities of preterm births among neighbourhoods. In the City of Windsor and Essex County, it seems that maternal behavioural, psycho-social and health risks may play more important roles than the maternal SES risk in the formation of preterm births. Although the spatial clustering of preterm births still exists after controlling personal risk factors, investigation of maternal behavioural, psycho-social and health risks in the City of Windsor and Essex County may keep to reduce some of the inequalities of preterm births among the health regions. Specifically, the statistical significance of the higher dependence on poor maternal health for preterm births suggests an accumulation of health impacts in this region, which need to be further addressed to identify the reasons.

Similarly, Figures G.13 and G.14 show distributions of spatially correlated random effects and the total random effects of the maternal health behaviour factor on LBW births. The fixed impact of behavioural risks is 0.1351 and the random effect ranges from -0.185 to 0.151. However the spatially correlated effect only contributes of 6% of the total area level variation. The random effect seems to be spatially randomly distributed.

For spatial impacts of the maternal SES risk factor (Factor 2) shown in Figures G.15 and G.16, 10% of the variation is contributed by the spatially correlated term. Similarly to preterm births, the



maternal low SES factor has a relatively lower impact on LBW births in the City of Windsor and Essex County.

For spatial impacts of the psycho-social risk factor (Factor 3) shown in Figures G.17 and G.18, Halton Region shows a relatively higher impact on LBW births compared to the other two regions. The spatially correlated random effect contributes 95% of the total variation. The fixed impact is 0.1143 and the random effect ranges from -0.25 to 0.23.

Figures G.19 and G.20 show that maternal health risks (Factor 4) may have a larger impact on LBW in rural areas than in the cities. The whole City of Windsor and Essex County show a relatively higher impact of the mothers' health risks than the other two health regions although large variations exist in the City of Windsor. The total effect is 0.2121 and the random effect ranges from -0.21 to 0.34. The spatially correlated effect contributes 22% of the total variation.

Overall, for the higher LBW risks in Windsor and Essex County, the mother's health situation and irresponsible health behaviours may play more important roles than the other two risk factors in determining LBW. These are shared with the above identified higher personal risks of preterm births. This means that mothers' health challenges, and adverse health behaviours, such as smoking, drug use, no prenatal care, and other careless behaviour caused by low education, financial difficulties, single parenthood, and family violence, may have more elevated impacts on LBW and preterm births in the Windsor and Essex County Health Region than in the other two regions. Further analyses may be needed to examine these potential accumulated impacts once the environmental causes of adverse birth outcome clusters are identified.

The credible intervals given by the Bayesian posterior distributions show that the 95% limits for most of the random effects of personal risk factors include 0, which means that there is a possibility that the higher or lower values of personal risks are noise from a random distribution. However, a confidence limit provides only information about how precise a given numerical value obtained from a sample is. It is greatly affected by sample size. The large ranges of variation for almost all personal risk factors compared with their fixed effects indicate that the spatial variations cannot be ignored. The high proportions of spatially correlated random effects for many of these variations further indicate the non-ignorable spatial influences. These results provide at least directions for future qualitative and quantitative research to address local clustering of adverse birth outcomes and suggest intervention directions for local authorities and stakeholders to address adverse birth outcomes according to local conditions.

One of the most plausible causes for the clustering of LBW and preterm births in the City of Windsor and Essex County is air pollution generated in the United States. Windsor is downwind from several strong air polluters across the Detroit River, pollution from coal-fired power plants in the

Ohio valley, or emissions from diesel trucks crossing the US-Canada border. The air quality in Windsor is regularly the worst in Canada. Many health problems, including respiratory illnesses and cancer, are more prevalent in Windsor than other places in Canada (Gilbertson and Brophy, 2001). The hospital admission rates are also significantly higher for cardiovascular and respiratory admissions in Windsor than in London, Ontario, indicating potential pollution-related risks (Fung, et al., 2007).

The identified clusters of preterm and LBW births in Figures 5.11 and 5.12 were further plotted against transportation networks to examine whether or not there are potential impacts of traffic generated air pollution on adverse birth outcomes. The maps are presented in Figures 5.19 and 5.20 for preterm and LBW birth clusters respectively. Figure 5.19 shows that, for preterm births, although the highest clustered areas contain the major corridors from the two Canada-US entrances (the Ambassador Bridge and the Detroit-Windsor Tunnel) to the west end of Highway 401, there are higher risks also along with the Detroit River, indicating potential impacts of pollution from across the Detroit River combined with from the high traffic volume. Figure 5.20 shows that the highest risks for LBW births are along the major transportation routes from the two Canadian-US entrances to the west end of Highway 401, indicating potentially somewhat higher impacts of traffic pollution on LBW.

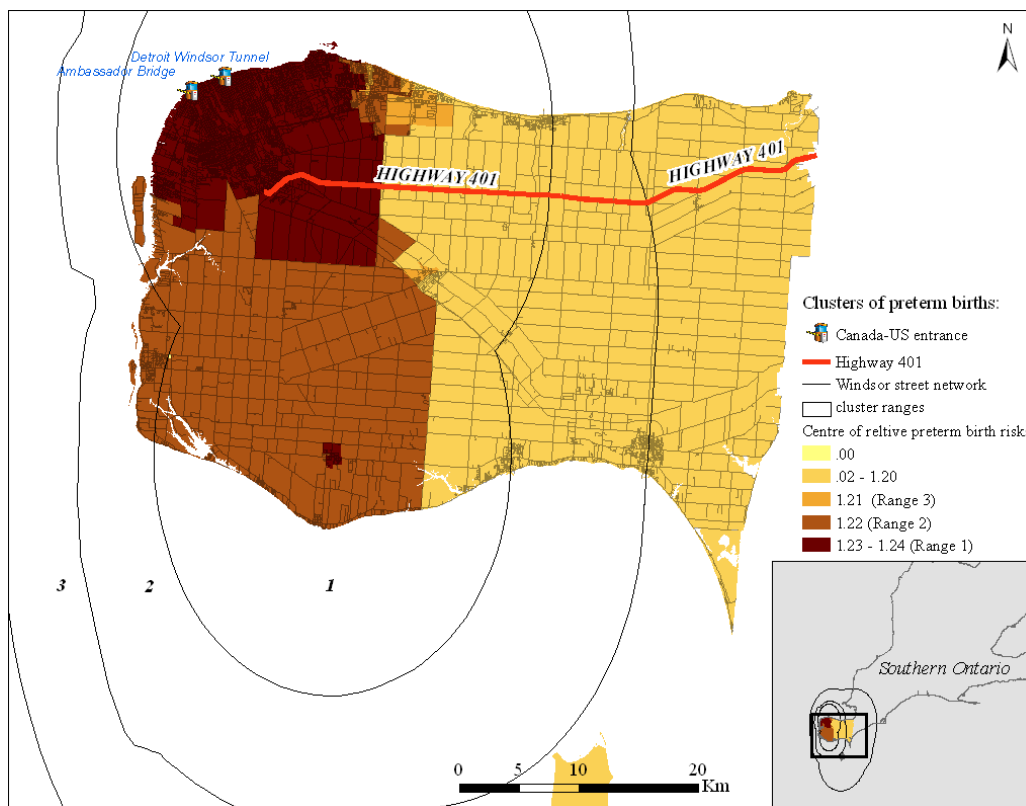


Figure 5.19: Traffic impacts on preterm birth clusters

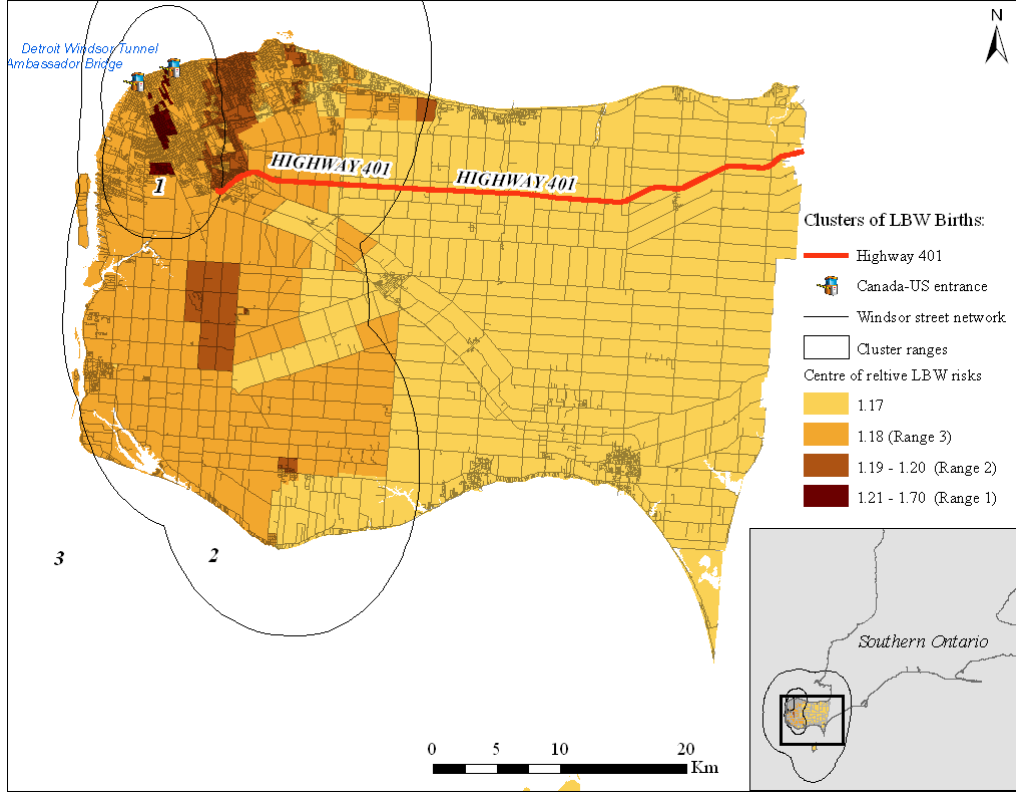


Figure 5.20: Traffic impacts on LBW birth clusters

To find out whether there is still an elevated risk in Windsor after controlling for personal risk factors and identified socio-economic, ethnic, and psycho-social risks, a three level model was constructed for preterm births and LBW births respectively to measure the remaining spatial random effects at the municipal level. The models are listed as in Equations (5.8) and (5.9).

*Preterm birth ~ binary ( $p_{ijk}$ )*

$$\text{Level 1 (personal): } \text{logit}(p_{ijk}) = \beta_{0jk} + \beta_1 \text{AGE19}_{ijk} + \beta_2 \text{AGE36}_{ijk} + \beta_3 \text{FEMALE}_{ijk} + \beta_4 \text{MLTIBIRTH}_{ijk} + \beta_5 \text{factor1}_{ijk} + \beta_6 \text{factor2}_{ijk} + \beta_7 \text{factor3}_{ijk} + \beta_8 \text{factor4}_{ijk}$$

$$\text{Level 2 (neighbourhood): } \beta_{0jk} = \gamma_{00} + \gamma_{01} \text{LOW\_SES}_{jk} + \gamma_{02} \text{CHC}_{jk} + \gamma_{03} \text{IV}_{jk} + v_{0jk}$$

$$\text{Level 3 (municipality): } \gamma_{00k} = v m_k + u m_k \quad (5.8)$$

*LBW ~ binary ( $p_{ijk}$ )*

$$\text{Level 1 (personal): } \text{logit}(p_{ijk}) = \beta_{0jk} + \beta_1 \text{AGE19}_{ijk} + \beta_2 \text{AGE36}_{ijk} + \beta_3 \text{FEMALE}_{ijk} + \beta_4 \text{MLTIBIRTH}_{ijk} + \beta_5 \text{PRETERM\_BIRTH}_{ijk} + \beta_6 \text{factor1}_{ijk} + \beta_7 \text{factor2}_{ijk} + \beta_8 \text{factor3}_{ijk} + \beta_9 \text{factor4}_{ijk}$$

$$\text{Level 2 (neighbourhood): } \beta_{0jk} = \gamma_{00} + \gamma_{01} \text{LOW\_SES}_{jk} + \gamma_{02} \text{EU}_{jk} + \gamma_{03} \text{VIS\_MINO}_{jk} + v_{0jk}$$

$$\text{Level 3 (municipality): } \gamma_{00k} = v m_k + u m_k. \quad (5.9)$$

Selected neighbourhood-level risks were included in these models to examine different aspects

of identified social and environmental determinants. The selection of neighbourhood-level variables was based on representing maximally the overall identified neighbourhood risks while avoiding any collinearity issues caused by correlations among the neighbourhood-level variables. For preterm births, neighbourhood-level socio-economic status, chronic health conditions, and insufficient vegetable intakes were selected to represent neighbourhood risk. For LBW births, neighbourhood-level low socio-economic status, emotional unhappiness, and visible minorities were selected. The term  $um_k$  is the spatially correlated random effect at the municipality level and  $vm_k$  is the uncorrelated random effect. As discussed earlier, the intrinsic CAR model with a contiguity neighbourhood structure depicts the neighbourhood-level spatial dependencies. At the municipality level, given the large sizes of municipalities, spatial dependencies may only exist among adjacent municipalities. The contiguity neighbourhood structure is good enough to describe this level of spatial dependency. The resulting total effects at the municipal level,  $vm_k + um_k$ , are mapped in Figures 5.21 and 5.22 for preterm and LBW births respectively.

Figure 5.21 shows that after controlling identified risks at both the personal and neighbourhood levels, preterm birth rates are still significantly high in the City of Windsor and the Town of Tecumseh. The LBW birth rate is high in City of Windsor but not significantly high for the Bayesian statistic, indicating potential mediational effects of identified personal and neighbourhood level risks. Nevertheless, the remaining higher risks of adverse birth outcomes suggest that further analysis is required to identify the impacts of potential environmental risks, especially air pollution, on adverse birth outcomes. Since air pollution data were not obtained for this thesis, it is impossible to examine the association between air pollution and adverse birth outcomes. However, further environmental risk research is clearly important.

Based on the above local spatial analyses, “hot spots” of adverse birth outcomes and significant variation of personal risks have been identified. After controlling for identified risk factors, spatial variation still exists, indicating potential influences of unidentified risks. Thus, for Hypothesis 5, the null hypothesis,  $H5_0$ , that risks of adverse birth outcomes, at both the personal and neighbourhood levels, have only universal or global impacts on adverse birth outcomes is rejected and the alternative hypothesis,  $H5_1$ , that local “hot spots” of adverse birth outcomes exist due to the impacts of surrounding neighbourhoods and personal and neighbourhood level risks affect adverse birth outcomes differently from neighbourhood to neighbourhood is supported.

The above five hypothesis tests confirm the impacts of space and place on adverse birth outcomes and suggest potential pathways of socio-economic, ethnic, living condition, and psycho-social risks on adverse birth outcomes. Based on these test results, recommendations on health interventions to improve birth outcomes are now suggested.

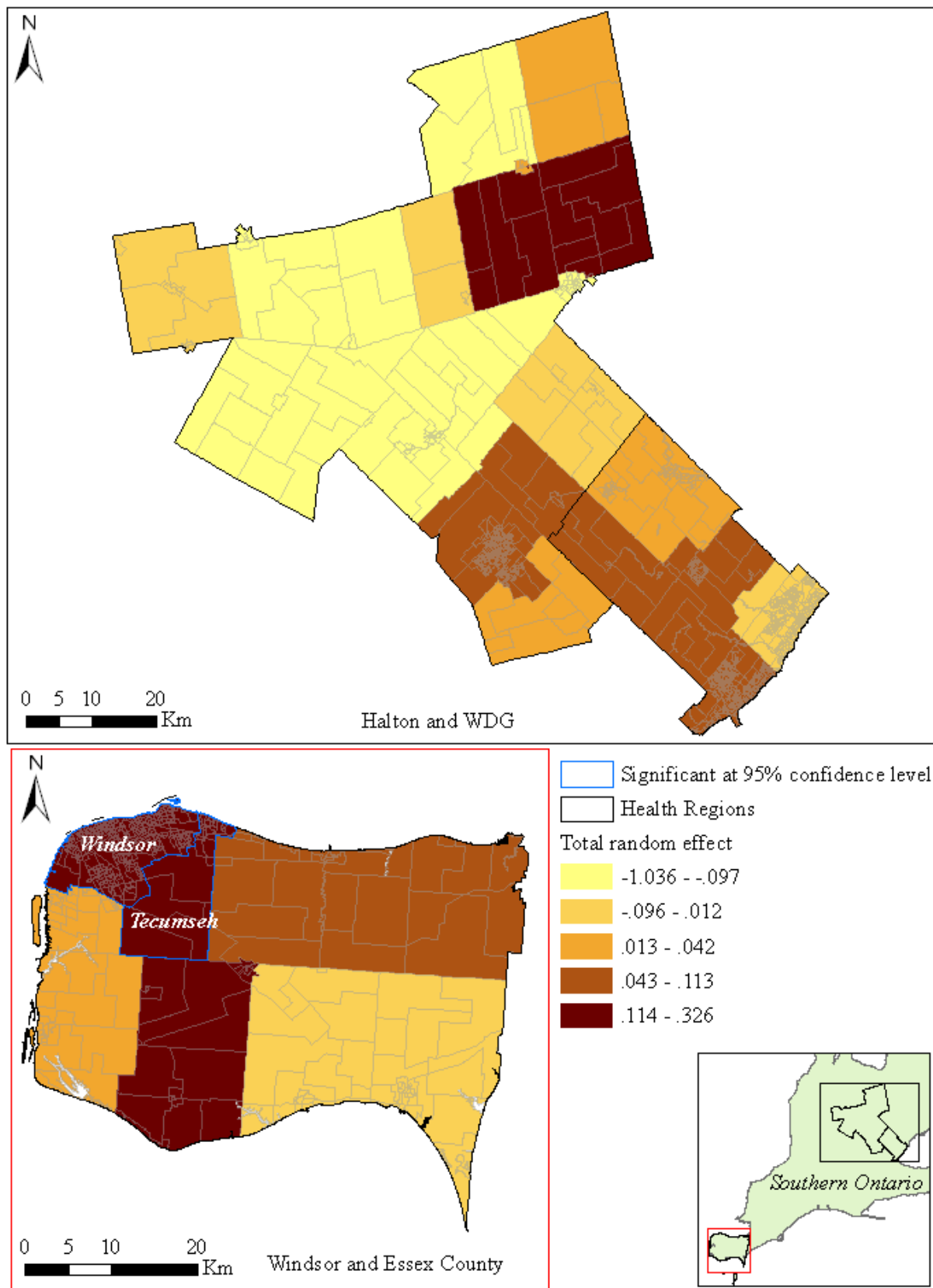


Figure 5.21: Total municipality level random effects of preterm births after controlling identified risks

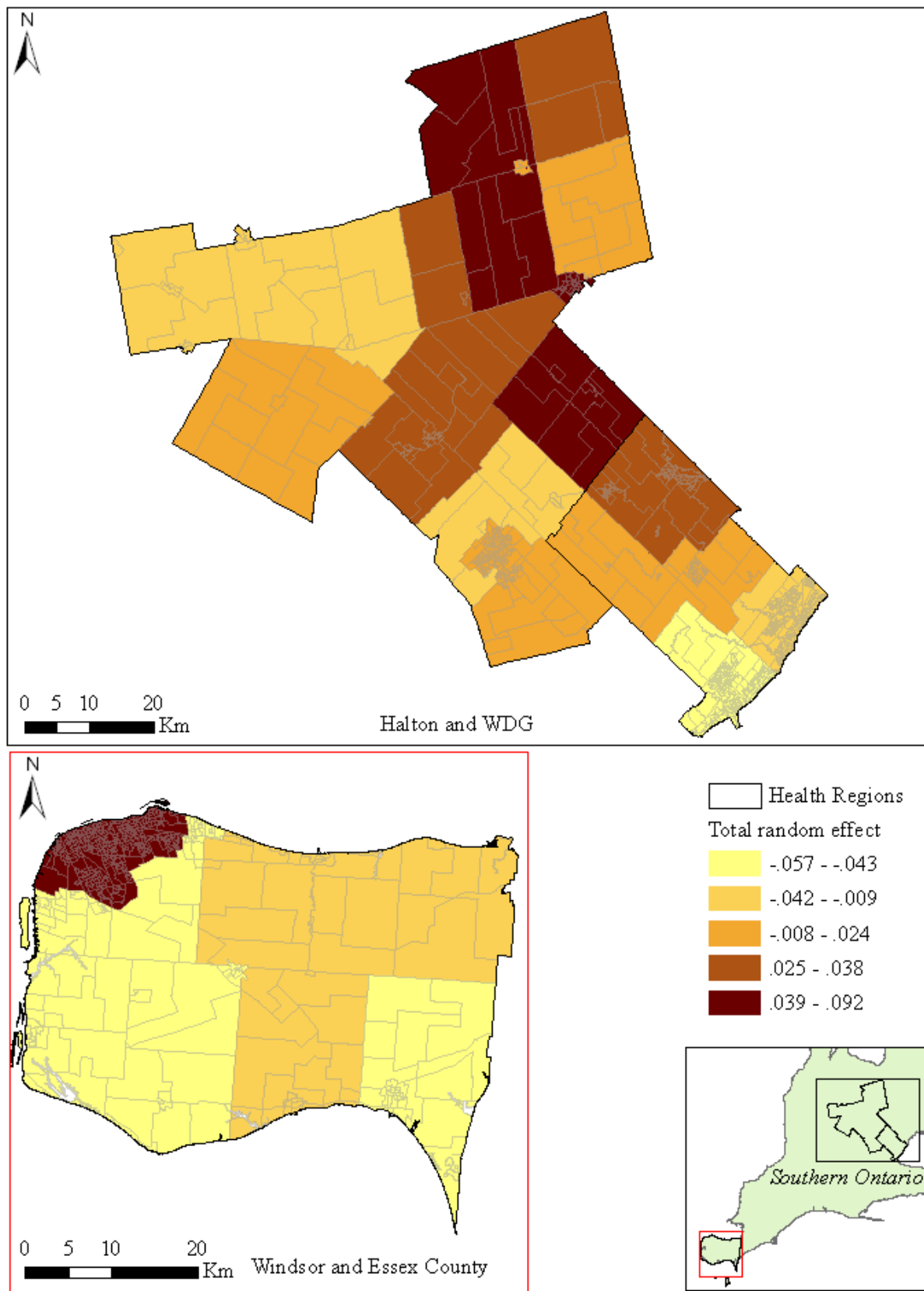


Figure 5.22: Total municipality level random effects of LBW births after controlling identified risks

## **5.4 Discussion and recommendations on community-based intervention and health planning**

The result of testing Hypothesis 1 shows that adverse birth outcomes in the study area have relatively small spatial variations, but are highly autocorrelated. This suggests that a universal neighbourhood-level intervention may not be so effective in reducing the overall adverse birth outcomes in the study region. However, a focused intervention on high risk neighbourhoods should be able to reduce the inequalities of adverse birth outcomes. Different spatial scales of adverse birth outcome clusters also provide evidence supporting different levels of intervention, from neighbourhoods to municipalities to health regions.

It is easy to understand the small spatial variation of adverse birth outcomes in the study region given the universal availability of health insurance in Canada, the various aspects of types of social welfare that are provided, and the relatively weak social discrimination against minority groups in Canada. However, high clusters of adverse birth outcomes are still present and are associated with various socio-economic, environmental, ethnic, and psycho-social risks. Despite moderate spatial variations, the associations between adverse birth outcomes and neighbourhood-level risks identified in this research are valid. For personal and neighbourhood-level risks that were already identified by other research, the findings are mostly consistent with the literature and they may be generally used to explain adverse birth outcome risks everywhere in socially advantaged societies.

If spatial inequalities of adverse birth outcomes are seen as the results of uneven spatial distributions of neighbourhood-level risks, the relatively evenly distributed neighbourhood-level characteristics and associated adverse birth outcomes in the study region may not suggest that the situation is evenly good. Some neighbourhood-level characteristics may also be evenly poor. For example, the variable of average sense of not belonging to local community is measured on a scale of 1 to 4, representing strongest to weakest senses of belonging to a local community. The actual range obtained from the data was 1.8 to 2.9 in the study region. None of the local neighbourhoods show strong community ties. As these community ties show significant association with LBW births and barely insignificant association ( $p=0.14$ ) with preterm births, the improvement of neighbourhood social ties or social capital may help to improve the overall birth outcomes. In addition, the variable of self-perceived stress does not show spatial dependence at the neighbourhood level (Table 5.1), indicating that personal stresses may be caused by other events and issues in life. This proves further the lack of neighbourhood mechanisms in the study region to release personal stresses and buffer their health impacts.

As identified earlier, the non-attendance at prenatal classes has no influence on having adverse birth outcomes, due probably to insufficient knowledge gained in these classes about adverse birth

outcome prevention or attending prenatal classes at a late stage of pregnancy. This suggests not only the inclusion of more adverse birth outcome prevention knowledge in prenatal education programs, but also the necessity of early stage intervention and education, even before pregnancies. This may be done through a community-based health promotion program. Knowledge of health promotion and adverse birth outcome prevention can be passed on to local residents through this program. Following the information gained, local residents can maintain their reproductive health in an organized and informative way by reducing unexpected pregnancies, avoiding unhealthy behaviours before and during pregnancy, and being well-prepared for their parenthood.

The factor analysis results (Table 5.11) show that the high risks for mentally challenged mothers or mothers with a family history of genetic health challenges to have LBW births are diminished for the derived genetic factor. This suggests that there may not be direct genetic influences for these mothers to have LBW births. It is more likely that the associated low socio-economic status of these mothers is the underlying reason causing LBW births. It can be shown in Table 5.11 that the variable of a mentally challenged mother also contributed to the SES factor. Thus, social and financial supports to genetic health challenged mothers may help to reduce their negative impacts on the LBW birth rates.

The confirmed compositional and contextual impacts of adverse birth outcomes through testing of Hypothesis 2 suggests that investigations of both personal and neighbourhood-level risks are necessary to reduce the negative impacts of neighbourhood-level risks. Neighbourhood-level socio-economic status along with related physical living conditions and food and vegetable deficiencies are the most robust neighbourhood-level contextual risks affecting adverse birth outcomes. These factors should be directly addressed.

In Ontario, personal food security is provided by a Special Diet supplement (up to \$250 per month) for poor people who are required to buy special food for medical reasons. While this may increase their ability to consume more food, low income people may still not eat properly if healthy and nutritious food cannot be accessed easily within or nearby local communities. The positive association between neighbourhood-level food and vegetable deficiencies and adverse birth outcomes indicates potential impacts of food supply on adverse birth outcomes and suggests the necessity for local food supply interventions. Providing healthy food at local grocery stores in low socio-economic status areas and supplying nutritious food and vegetables directly to low income pregnant women may reduce the incidence of adverse birth outcomes.

The pathways resulting from the analysis of Hypotheses 3 and 4 show the complexity of neighbourhood-level determinants on adverse birth outcomes. Health promotion programs can benefit from an understanding of these complex structures of influences. For example, the influence of



neighbourhood-level ethnic composition on adverse birth outcomes may not be simply understood by the low socio-economic status of the minority groups. The research findings identified potential positive impacts of minority group cultures on maternal health behaviours and potential negative impacts of the social isolation of low-income ethnic majority groups on birth outcomes. Hence, health promotion programs can be developed that seek to promote group cultures and increase the level of social integration between minority and majority groups.

Given the currently prevailing globalization and social polarization processes that are in progress worldwide, the increasing social and economic gaps between the rich and the poor are not going to be reversed in a short period of time. Thus, it may not be feasible to reduce neighbourhood socio-economic inequalities through the reduction of personal socio-economic inequalities. In fact, the increasing polarization of socio-economic classes has led to the division and segregation of residential areas and has allowed socially isolated and economically deprived communities to develop. To reduce the negative impacts of socio-economically disadvantaged communities or to break up the concentration of poverty in inner-city or inner-suburban neighbourhoods is one of the greatest challenges confronting in the metropolitan areas of developed countries. Poor health is also one of the many challenges facing the local residents of such communities. This research provides further evidence that adverse birth outcomes are associated with such communities.

Initiatives such as mixed-housing types through direct developer subsidies, as well as mixed-tenure or income mixing residential development based on new urbanism principles have been proposed by urban planners to address residential segregation (Congress for the New Urbanism, 1999). These initiatives try to promote diverse and liveable neighbourhoods with a large variety of housing types and land uses to serve a wide range of households with different sizes, ages, incomes and cultures. This has changed the way that public housing is constructed in many communities. In fact, mixing the population of low-income groups with better-off groups in residential development has become a common policy in many Western countries (Smats and Salman, 2008). The concentration of public housing projects in one or several adjacent neighbourhoods in the past has led to the concentration of the poor and has created many associated social and health problems. Through reintroducing new mixed-type or mixed income housing in socially disadvantaged communities or relocating public housing into mixed-type communities, a mixed-income environment with greater economic and social diversity may be created to reconnect low-income residents to their surrounding neighbourhoods and thereby enhance their chances to re-enter into mainstream society.

The purpose of such development by mixing different social classes, especially low- and medium-income classes, is to build “communities” in order to reshape environments, opportunities, and social arrangements that can promote healthy neighbourhood life. However, earlier projects

showed complex results (Musterd and Andersson, 2005; Joseph and Chaskin, 2010). The quality of physical living conditions for low-income families was improved by living in a new, clean, well-managed development in the midst of a revitalised neighbourhood. The opportunity to observe their middle-class neighbours also led to changes in aspirations and behaviours of the low-income population. On the other hand, other expectations including expanding social networks to increase access to information, resources, and social opportunities and increasing informal social control and collective efficacy were not realized due to potential social barriers between different social groups. This is because the strategy of physical mixing of different social classes alone considered only partially the reasons for the formation of ghetto culture and behaviours. The true barriers to social and economic mobility, including access to employment, were not addressed (Smith, 2006). Potential social integration may not be achieved if no further intervention is involved. Hence, the social mixing strategy alone cannot solve problems caused by social inequalities.

Social mixing is not without its advantages for the health of local residents. The adverse birth outcome analysis results in this thesis suggest that housing mix or social mix may be beneficial for addressing adverse birth outcomes in terms of improved physical living conditions and health behaviours of the low income groups. However, the earlier experiences discussed above indicate that a mixing strategy needs to be combined with other community-oriented strategies that can bring people of diverse ages, ethnic backgrounds and incomes into daily interaction and strengthen the bonds of local residents with their communities. While daily interactions may be partially promoted through better mixed land-use designs, it is also important to increase the social capital of local communities through community programs to form neighbourhoods into coherent units.

The research findings provide implications for incentives and focal points for community-oriented health promotion programs. Public health authorities may act as initiators and facilitators to initiate and support such programs in neighbourhoods identified with high risks of adverse birth outcomes. Local residents and stakeholders should make up the major body for discussions and decision making over their solutions. Based on the identified social and environmental risks and the pathways suggested in the testing of Hypotheses 3 and 4, local residents and stakeholders may know not only what actions to take, but also how to implement programs. The analysis of local variations of personal-level risks through the testing of Hypothesis 5 also provides evidence on what risk/or risks are more influential based on the local context.

If the preterm birth rate is high in a neighbourhood, local social and environmental risks may be first sought among socio-economic classes, ethnic groups, different physical living conditions, and areas with existence of environmental pollution. Once potential risks are identified, multi-component interventions can take place, focusing on identified contextual risks and personal risks that these

neighbourhood-level risks may lead to. If it is found that the proportion of Aboriginal population is high in a neighbourhood, the focus should be on improving the socio-economic conditions of the mothers and reducing the adverse impacts caused by their low SES. If other minority ethnic groups are concentrated in the neighbourhood, the focus should be on the negative impacts of a mother's SES and health conditions, as well as the connection between majority and minority ethnic groups. This may be particularly meaningful for low SES majority groups in such a neighbourhood. If health services or food or vegetable supplies are shown to be insufficient within the neighbourhood, collective efforts can be generated toward the improvement of such services. Even though most personal-level risks are found to have potential impacts on the neighbourhood, the local spatial analysis results also suggest which one/ones should receive attention first.

In addition, a community-based health promotion strategy would have to accept that health disparities from socioeconomic differences are best understood by the people who experience the greatest ill-health burden. They therefore constitute a major resource for guiding examination of the circumstances and events that influence health outcomes. Based on the information provided above, local residents and groups that are most affected by adverse birth outcomes may better understand and identify clearly what may affect them the most. Unlike multi-component programs discussed in the literature, such community-oriented interventions are more focused on identified local risk factors and may have influences on all local residents, not just on pregnant women. Couples who plan to have babies may potentially benefit from such programs as well.

Through increasing participation of local residents to foster their common health goals, these programs may also increase the social capital of neighbourhoods. Since people from similar social backgrounds may be more attracted to each other as neighbours, families of different social classes may not be willing to interact with each other for leisure. However, they may be willing to work together to realize their common goals. Such health promotion programs may not only enhance a neighbourhood's collective capacity for action to gain control over their health conditions and other neighbourhood affairs, but also lead to an active sense of engagement and bring local residents together as a cohesive unit. This will increase social interactions and information sharing between different classes. The increased feelings of control over residents' life events and health conditions, and their increased sense of belonging to local communities may also have a beneficial effect for releasing stresses and other psycho-social tensions, and eventually have buffering effects on adverse birth outcomes.

However, practically speaking, without knowing the benefits and effects of such community-oriented programs, it is not possible to initiate such programs in all communities. Based on the analysis results presented in this chapter, a bottom-up strategy may be more feasible starting from the

most vulnerable communities as pilot projects. Once local public participation has been mobilized and these projects have shown positive impacts on improving the health of local residents, new projects may be started at other neighbourhoods based on the experiences gained from previous projects. Along with the increasing engagement of neighbourhoods, public participation may be fostered at higher levels which may involve representatives from different neighborhoods and different interest groups. Health problems within different contexts may be identified and consensus and solutions may be made on higher level affecting factors. By stimulating the awareness of local residents or groups, these groups can then press from the bottom-up for fairness on general thinking, such as welfare systems, taxes, minimum wage protection, medical care systems, public housing provisions and other social security systems. This lobbying may then be aggregated and mobilized into collective efforts to improve population health, such as birth outcomes in this case, and improve people's quality of life.

To reach this goal, temporary or research-based projects on community-based health promotion programs may be of limited benefit. Such programs should be relatively stable both financially and structurally and have a mid-to-longer term horizon. In Ontario, the promotion of community engagement in health service plans and the creation of the Local Health Integration Networks (LHINs) provide the opportunity to support community-oriented programs. With the devolution of responsibility to the LHINs, a greater emphasis has already been given to finding ways to involve local community members, health care organizations and other partners in discussions and decision-making about the delivery of health and social services. It is possible that financial support may be sought through a local LHIN to support these programs. With this support, trained personnel may be assigned by the health authority to a neighbourhood health promotion program as a coordinator to initiate and manage local health affairs and connect between health authorities and local stakeholders. The person may work in a local community centre and also be a consultant for daily health-related information requirements (such as the prevention of adverse birth outcomes). Local representatives and stakeholders can be gathered to discuss and make decisions on health issues. A local board may also be established for governance on a voluntary basis consisting of local elected members to represent interests of different stakeholders. Through this setting, health-related information can be shared, public participation can be stimulated, public opinions can be collected, decisions can be made on the consensus of different interest groups of the local residents, and negotiations can be initiated between health authorities and local residents over issues such as health care provision, health policies and plans to affect final decision-making. Through the empowerment of local citizens (including the most vulnerable populations), health care may eventually be delivered based on public health need.

Thus, in addition to the currently prevailing personal-oriented health intervention approach, community oriented health promotion programs should be introduced to address the various social and environmental risks related to adverse birth outcomes. The combination of these two approaches may be able to reduce the overall occurrence of adverse birth outcomes and, at the same time, reduce their inequalities among neighbourhoods and among social classes.

## **5.5 Summary**

In this chapter, the proposed five hypotheses were tested through an empirical study conducted in three health regions in Ontario, Canada, namely the WDG, Halton, and WEC health units. Various statistical models were constructed to test the direct and indirect roles that space and place play on the courses and outcomes of pregnancy and on the social and environmental determinants of these outcomes. Pathways by which various neighbourhood-level risks, including socio-economic, ethnic, living conditions, psycho-social, and health service risks, influence personal and behavioural risks were discussed and identified. “Hot spots” of high incidences of LBW and preterm births were located and possible related causes were discussed. Urban planning and community-based health promotion programs to reduce spatial and social inequalities of adverse birth outcomes were eventually suggested.

## **Chapter 6**

### **Conclusion**

This thesis has thoroughly analysed the social and spatial determinants of adverse birth outcomes, especially at the neighbourhood or community level. However, the constraint of data unavailability at the personal and neighbourhood levels limited the ability of the pathway analysis to generate a more complete picture of the risks of adverse birth outcomes. Based on the knowledge and theoretical assumptions identified in previous research, the thesis goal and objectives were achieved using various statistical and spatial analyses to test five research hypotheses. In this chapter, the major contributions of this thesis are summarized. The limitations of this work and future research directions to improve adverse birth outcome analysis are discussed.

#### **6.1 Thesis achievements**

##### **6.1.1 Theoretical achievements**

Although many different theoretical explanations of the pathways of health inequalities have been suggested, the research literature of adverse birth outcomes shows that no research has yet tested statistically the pathways of adverse birth outcomes and their inequalities. The social and environmental determinants of adverse birth outcomes were only relatively recently analyzed by researchers and are not comprehensively understood. Aimed at improving this literature, this thesis has provided analyses of the general pathways of commonly identified neighbourhood-level risks of adverse birth outcomes. Rather than analyzing the pathways of a single risk in much detail and quantifying their influences, the research examined the overall interrelations between different aspects of neighbourhood-level and personal-level risks in their impacts on adverse birth outcomes. Statistical testing was used mainly to establish the relations and to confirm whether or not their hypothesized nature, based in the literature, could be statistically supported. As a first attempt to seek the pathways of neighbourhood-level risks to adverse birth outcomes, the identification of effective pathway patterns and major neighbourhood-person interrelations is more important for further understanding and analyses than quantifying the magnitude of each effect, although this was also calculated by the models. Once pathways are generally determined, more quantitative measures can be pursued on more detailed data to seek specifically the magnitudes of different pathways of a certain neighbourhood-level risk. This may provide direct evidence for health intervention and resource allocation.

In the thesis, the general pathways of socio-economic, ethnic, living conditions and psycho-social risks to adverse birth outcomes through personal-level risks were identified using mediational analysis. Three hypothesized types of pathway models, namely the behavioural model, the materialist model and the psycho-social model, were tested at the neighbourhood level. Both the materialist and psycho-social models were supported by the data, although the direct material impact of living conditions on mother's health is not as well supported as the socio-economic, behavioural and psycho-social pathways, and neighbourhood-level psycho-social factors were found to play a relatively more important role on the onset of LBW (or IUGR) than preterm births.

The direct behavioural explanation of personal predisposition to health behaviours and to adverse birth outcomes was not supported. Positive cultural impacts of group influences in ethnic groups on reducing adverse birth outcomes were identified. The associated low socio-economic status of minority ethnic groups, especially for Aboriginal populations, was found to play an important role on the onset of adverse birth outcomes. The research also found that non-ethnic low-SES residents in neighbourhoods of high ethnic compositions may suffer more from accumulated social, cultural, and economic disadvantages than other ethnic and socio-economic groups. All these identified pathways set up potential directions for future research and interventions to address the inequalities of adverse birth outcomes.

In addition, efforts were made in the thesis to identify local spatial impacts of adverse birth outcomes. The open and evolving social system and the complexity of social factors determine that it is unlikely to be possible to identify all potential risks and their fixed universal impacts on adverse birth outcomes. Even if all potential risks could be identified, changing social situations may reveal new conditions and challenges that affect the identified associations. Given their complexity, it is also not feasible in practice to investigate all social and environmental risks, some of which are beyond the health sector domain. Thus, local spatial analyses were used to locate places of high adverse birth outcome incidences and to identify potential influences of local contexts on these high risks. This aspect of the research is new to the existing and large adverse birth outcome literature.

In the thesis, the "hot spots" of significantly high incidence of adverse birth outcomes were identified at different spatial scales through local analysis using the spatial scan statistic and multilevel spatial random effect models. The potential associated risks were identified and discussed and future analyses were suggested. The spatial variation of personal-level risks were also obtained and mapped in the study region allowing the direct and indirect roles that space and place play on the courses and outcomes of pregnancy to be identified. Findings not only legitimate the focus of community-oriented interventions and point out focused locations for intervention, but also provide evidence on how to address the social and personal problems associated with these areas. Through

health surveillance and local environmental risk evaluation, health interventions can be conducted pertinently and effectively based on these results, focusing not only on high risk communities, but also on high influence risk factors.

### **6.1.2 Methodological improvements**

To achieve the goal of the thesis, various statistical means were utilized for the tests of the five research hypotheses. A series of procedures was developed to make possible a mediational analysis on multilevel binary outcome models so that the pathways of neighbourhood-level risks on binary adverse birth outcomes (LBW and preterm births) could be tested. Although the estimation is approximate, the improved procedures provide new thoughts and a means to realize this kind of analysis.

Attempts were also made in the thesis to construct various multilevel spatial models for testing the impacts of social and spatial determinants on adverse birth outcomes from different angles. Although it is already established that the use of multilevel models is necessary for social and environmental risk analyses, and multilevel model building has been suggested by various researchers, only a small proportion of studies have so far used multilevel settings for the analysis of adverse birth outcomes. The uses of multilevel models in the thesis not only helped to identify different impacts at different levels on adverse birth outcomes, but also made it possible to discover the between-level interrelations. The constructed models are valuable for adverse birth outcome analyses in other regions for purposes of both health surveillance and future research. Once data become available on unidentified aspects of social and environmental determinants, these models can be extended simply to accommodate new exposure variables.

The complexity and the interaction of individual, social and environmental risks determine that it is impossible to construct a single model to incorporate all of the potential risks for analysis. However, the models constructed in the thesis not only provide a practical means for the proposed analyses, but also bring threads of thought for spatial analysis of other health outcomes in similar settings. Through careful design and controlling for potential confounders via multilevel model building, the “true” impacts of a certain environmental risk on health and disease may be revealed.

The use of Bayesian hierarchical model building techniques made the multilevel models more flexible. In the thesis, spatial impacts were able to be modeled explicitly using a spatially structured random effect item in the Bayesian models. To reveal the underlying spatial dependency of adverse birth outcomes, different neighbourhood structures were constructed and compared. The most suitable neighbourhood structure was identified through modeled effect comparisons. Spatial variations after controlling for confounders at both the personal and neighbourhood levels were



captured effectively using the identified neighbourhood structure in Bayesian hierarchical models. The comparison procedure for neighbourhood structures can be extended to encompass seeking the best suited spatial weight matrix to depict local spatial dependency.

Since many of the identified personal level risks are correlated, they cannot be used in a single model to depict the general impact of personal-level risks. However, it is necessary to have a relatively general description of personal-level risks so that pathways of social and environmental risks toward different aspects of personal-level risks can be analyzed. To address this concern, a factor analysis was conducted on the identified personal-level risks for LBW and preterm births respectively. A principal component analysis was first undertaken, followed by an orthogonal rotation of the identified principal components using the varimax method. The extracted factors through this method were able to represent different aspects of personal-level risks. Given the independence between these factors, they can all be put into a single model to represent the overall impacts of personal-level risks. A further benefit of the identified factors is that the extracted “pure” genetic factor does not associate with preterm births, which provides some evidence against the assumption of genetic influences on adverse birth outcomes.

Another technical improvement in the thesis is the development of spatial interpolation procedures to obtain neighbourhood-level variables from secondary data, specifically, the Canadian Community Health Survey data. Through visual comparison and cross validation comparison on two commonly used interpolation methods, namely inverse distance weighting and Kriging, the best suitable Kriging method and corresponding parameters were identified for the spatial interpolation of CCHS data. The procedures developed are also expected to be useful for spatial interpolations of other health survey data to obtain community-level health-related indicators.

### **6.1.3 Practical implications**

Based on the results of the thesis analysis, health intervention policy improvements, especially community-based multi-component interventions, such as early stage prenatal health education and local provision of healthy food, were suggested in the last section of Chapter 5. The social and spatial analysis of determinants and pathways of adverse birth outcomes provide evidence on where and how to act to reduce the inequalities of adverse birth outcomes among neighbourhoods. Potential impacts of interventions beyond the domain of the health sector, such as social mixing through planning means, to remedy the social inequalities of adverse birth outcomes, were also discussed.

More importantly, an initiative for community-based health promotion was suggested. For health authorities and health practitioners, such a program means long-term health maintenance, early preparation, targeted risk avoidance and enhanced social care for not only the at-risk individual, but

also the regular residents. The realization of this initiative may move the current “sickness” care system to a true “health” care ideology. For local communities, especially for those which are socially and economically disadvantaged, this means putting public health back into the hands of the “public”. Instead of passively receiving money from the government for seeing a doctor when they are sick, local residents may actively care for their health through this program. They may also gain some power through collective efforts to argue or determine what needs to be done to improve local conditions for their health.

Although the association of socio-economic disadvantage and poor health is the most solid finding among all social determinants, and the correlation has been established over and over again for the last several decades, there is still no effective policy and health intervention developed in Canada or most socially advanced societies that focuses specifically on the reduction of health gaps between the rich and the poor. To narrow these gaps, community health initiatives have been proposed among various social groups, such as researchers, non-government organizations, feminists, and environmentalists. However, their passion for this approach has not necessarily translated into action. Health funds were often spent based on the decisions of people who have never met or interacted with a poor person. Priorities on government expenditures have usually been given for economic development and deficit reduction rather than poverty and inequality reduction. To move the latter agenda forward and turn the attention of the government around, a strong sense of purpose must be grounded in local knowledge and solid support.

Aiming at providing direct evidence on where and how to act, the research in this thesis has explored the pathways and local spatial impacts of adverse birth outcomes. This has been done in the expectation that the evidence can stimulate both local residents and the government to take action. Based on the evidence, a community-based health promotion program can be stimulated, starting from the most disadvantaged communities. By knowing what is happening around them, local residents will be purposively and actively recruited to work collectively in reducing the identified risks and protecting their health. This local knowledge in combination with collective efforts may increase the bargaining power of the “invisible” groups and act from the bottom up to bridge the gap between the government and local communities. Passion and informal actions may eventually translate into policy changes and by-laws. Sustainable funding may be put into the frontier of health care to support and consolidate community-based health promotion programs. Given the targets of social and environmental risks, these programs may also provide opportunities to work in collaboration with other sectors and aim at not only reducing adverse birth outcomes and improving the health of local residents, but also generally improving their quality of life.

## **6.2 Limitations and future research**

Although the goal of the research is achieved, there are still limitations and constraints that have impeded the current analyses. These issues should be addressed in future research.

### **6.2.1 Data constraints**

In this thesis personal-level data were obtained from only three public health regions in Ontario. These health regions are not geographically connected. This spatial disconnection may have had a potential impact on the spatial analysis. The isolation of the study areas may make the analysis results more subject to boundary effects, and the spatial autocorrelation between regional health units is not able to be examined. In addition, the study region contains mid-to-small sized cities and towns. The exclusion of large cities, such as Toronto, may lose some power to examine and discover more generally the associations between adverse birth outcomes and their social and spatial determinants. Socio-economic inequalities of health may be evident to a greater degree in large cities since they are more globally impacted and socio-economic and ethnic groups are more polarized or marginalized. The inclusion of more health regions, especially large metropolitan areas, in future research may not only help to enhance statistical power through increased sample sizes, but also help to build the interconnections among health regions and help to find in-depth correlations between health, space, and social risks.

In this thesis, neighbourhoods were assumed to coincide with census dissemination area boundaries. However, although CDA units are relatively homogenous, they are artificially divided for census purposes. Some neighbourhood characteristics, especially the social capital of a neighbourhood, may not be effectively captured by CDA surrogates. If naturally defined or behavioural neighbourhood boundaries could be identified, they should be used for future analysis.

More neighbourhood-level variables need to be collected, such as environmental pollution, health services, and social capital of the community, to explore further social and environment risks. As identified in Chapter 5, the cluster of adverse birth outcomes in Windsor is possibly due to the impacts of air pollution. The collection of such data and other potential environmental risks may help to clarify the exact reasons for the clusters.

The collected personal-level data were not without their disadvantages. Although a great deal of information is already collected in the ISCIS database, more personal-level risks need to be identified. Ethnic status, such as visible minority status and immigrant status, may be collected so that the ethnic impacts on adverse birth outcomes can be further clarified. Only 1.9% of all mothers were identified in the data as having low education status. This is much lower than the average Canadian rate reported by Statistics Canada (13% of Canadians aged 25 to 64 had not completed secondary school

in 2007), even allowing for the difference in age ranges. This probably is because the corresponding question in the screen tool only asks whether or not the mother has low education status. It should be made more specific in future information collection so that the impact of maternal low education on adverse birth outcomes can be clearly discovered.

Although not an easy job, below personal-level medical or etiological evidence should be collected focusing on the abnormalities of the mother, the foetus, and the placenta. Since higher-level risks will eventually affect one or a combination of these factors, knowing these factors can further assist identifying the causal pathways of social, environmental and personal risks. This will have significance for all primary, secondary and tertiary preventions. The data may be collected through a closer collaboration between PHUs and hospitals.

Data incompatibility may be another issue to be addressed. Personal-level data were collected from 2000 to 2008. A majority of the neighbourhood-level data were obtained from the 2001 and 2006 Census and from the 2001, 2003 and 2005 Canadian Community Health Surveys. Although the time periods are approximately the same, there are still some discrepancies, which may cause some inconsistency between the data collected and the real social and environmental status of the mother during pregnancy. Although the social and economic conditions of neighbourhoods were similar between 2001 and 2006 in the study area, 2001 census data were collected for the analysis based on the belief that mothers' and babies' health conditions are affected not only by the social and environmental conditions during pregnancy, but also by accumulated social and environmental impacts from the past. Nevertheless, if neighbourhood-level data can be collected annually and be consistent with the live birth data, the quality of social and spatial analyses may be improved and a longitudinal analysis may become possible to identify social and environmental impacts and the impacts of policy changes over time. The life-course approach discussed in Chapter 3, Section 3.1.4 may also be tested at the neighbourhood-level to reveal whether or not long time exposure to neighbourhood level risks may have accumulated impacts on adverse birth outcomes.

### **6.2.2 Technical constraints**

Even if data can be collected for all potential risks at all levels, given the complexity and the interrelations of different level risks, it is not possible to consider and analyze all potential risk factors in one model or in one project. Careful research design and proper statistical modeling are needed to take the complex interrelated phenomena apart and reveal the true impact of each risk.

Given the multilevel nature of the determinants of adverse birth outcomes, the analysis of their spatial variations at different spatial scales involves complex statistical modeling. It is difficult for a frequentist method to fit spatial models using a maximum likelihood or quasi-likelihood approach.

Bayesian statistics need to be involved to handle this complexity.

While the Bayesian models give a lot of flexibility in model building to measure complex social and environmental determinants of health, the calculation is time consuming. It took more than two days to run the random effect models using the WinBUGS tool on a quad-core processor on a personal computer. The WinBUGS tool is also very sensitive to the prior distribution. Although it would be nice to compare the intrinsic CAR model with the proper CAR model to discover better the spatial structure of adverse birth outcomes, the use of the proper CAR model as a prior to calculate the spatial dependency at neighbourhood level was not successful using the WinBUGS tool. The error messages provided by this tool (e.g. Trap 66: post condition violated) are not useful to identify potential problems and investigation could not reveal a solution or a cause for this. A more robust and user-friendly statistical tool for calculating Bayesian models, especially for calculating spatial random effect models, is yet to be developed.

The OpenBUGS project provides an excellent opportunity to improve further the current BUGS functions by sharing the source code. Through efforts of researchers from different fields, the tool is expected to be improved and become more robust and reliable. However, these developments have yet to become available.

Since the derived neighbourhood-level CCHS variables are from spatial interpolation, they unavoidably involve accountability issues. The Kriging interpolation method provides a beneficial outcome, namely the production of standard errors for predicted variables (Figure 5.8). Uncertainties in the interpolated values are described by these standard errors. It can be observed in Figure 5.8 that for locations with very few samples, the standard errors are higher, the confidence intervals are larger and the interpolated results are less accurate. These standard errors can be used directly in statistical modeling, such as Bayesian spatial hierarchical modeling, to handle data uncertainty issues and improve the reliability of analytical results using these derived variables. For example, the multilevel Bayesian models that involve CCHS variables in Chapter 5 can be simply modified for the neighbourhood-level model as:

$$\begin{aligned} \text{Level 2 (neighbourhood): } \beta_{0j} &= \gamma_{00} + \gamma_{01}NB\_VAR_j + u_{0j} \\ NB\_VAR_j &\sim \text{normal}(NB\_RISK_j, NB\_VARIANCE_j). \end{aligned}$$

Instead of using the neighbourhood-level risk factor,  $NB\_RISK_j$ , as data, it can be considered as a variable with a distribution (such as a normal distribution) and the Kriging standard errors can be used to calculate the variances,  $NB\_VARIANCE_j$ . Data uncertainty issues can then be effectively handled and the analysis results will be more reliable. In addition to the stochastic standard errors obtained from Kriging interpolation, an alternative for the production of standard errors in

neighbourhood-level predictors is the use of small area methods which take account of spatial dependence (Pfeffermann, 2002). Similar models can then be constructed using the produced standard errors.

In addition to the models already constructed in this thesis, more multilevel models need to be sought to analyze upstream higher-level impacts, such as city or regional level policies, on adverse birth outcomes. In conjunction with further medical discoveries, the downstream pathways that go below personal-level risks also need to be studied to reveal the actual causal relationships between social and spatial determinants and adverse birth outcomes. As discussed earlier, longitudinal studies need to be conducted as well to find out not only space but also time influences on adverse birth outcomes.

The open and complex nature of social systems and the complexity of social and environmental risks mean that health geographical research on adverse birth outcomes is a long-term ever-changing task. Despite all the difficulties faced by researchers, based on plausible theoretical explanations, by careful design of cross-sectional, longitudinal, multilevel and community-based studies, it will be possible to take the causal nexus apart and identify the operating mechanisms of social and spatial risks to reduce the occurrences and inequalities of adverse birth outcomes. Although theoretical and technical progress has been made through this thesis, a great deal of work remains to be done.

## Appendix A

### Description of variable names

<i>VARIABLE NAME</i>	<i>DESCRIPTION</i>
<b>AGE19</b>	age of mother less than or equal to 19 at delivery
<b>AGE36</b>	age of mother equal or greater than 36 at delivery
<b>LANGNOTEN</b>	mother's language is not English
<b>FEMALE</b>	baby's sex is female
<b>PRETERM</b>	preterm birth
<b>HEALTHCLNG</b>	health challenges of the mother
<b>INFECTIONS</b>	mother's infection during pregnancy
<b>DRUGS</b>	drug and alcohol uses during pregnancy
<b>GENHLTHCLNG</b>	Family history of genetic health challenges
<b>SINGLEPARNT</b>	single parent
<b>NOSOCIOSPT</b>	no social support
<b>FINADIFC</b>	financial difficulty
<b>NOPRENCARE</b>	no prenatal care
<b>SCHIZMOM</b>	evidence of schizophrenia of mother
<b>SCHIZDAD</b>	evidence of schizophrenia of father
<b>LBW</b>	low birth weight
<b>POSTDEPRES</b>	post delivery depression
<b>MENTALMOM</b>	mentally challenged mother
<b>MENTALDAD</b>	mentally challenged father
<b>MARIDISTR</b>	marital distress of the mother
<b>LOWEDU</b>	low education of the mother
<b>FAMILYVIO</b>	family violence
<b>NOPRECLASS</b>	not attending prenatal classes
<b>STRSDELIVY</b>	stress related to delivery
<b>SMOKING</b>	maternal smoking
<b>UNDERNUTRI</b>	maternal under nutrition
<b>MULTIBIRTH</b>	multiple births
<b>NO_CITIZEN</b>	Percentage of none Canadian citizen
<b>IMMIGRANTS</b>	Percentage of immigrants
<b>VIS_MINO</b>	Percentage of visible minority
<b>ABORIGINAL</b>	Percentage of aboriginal people
<b>DWL_VAL</b>	Average value of dwelling

<b>UNEMPLOY</b>	Percentage of unemployment
<b>NONOFFLANG</b>	Percentage of none official language speaking population
<b>MOVERS</b>	Percentage of movers in the recent 5 years
<b>TOT_INC</b>	Average total income
<b>EMPLOY_INC</b>	Average employment income
<b>FAMILY_INC</b>	Average family income
<b>HOUSE_INC</b>	Average household income
<b>LOW_INCOME</b>	Percentage of low income population
<b>PERROOM</b>	Average person per room
<b>PERBEDROOM</b>	Average person per bedroom
<b>RENT_RATE</b>	Average rental rate
<b>MAJHOSREPA</b>	Percentage of major repaired houses
<b>LOW_EDU</b>	Percentage of low educated population
<b>LOW_SES</b>	Percentage of Low socio-economic status (composed by LOW_INCOME, UNEMPLOY, and LOW_INCOME)
<b>CHC</b>	Average chronic health conditions, range 0-1 (0-no chronic condition, 1-has chronic condition)
<b>RD</b>	Percentage of regular drinker (range: 0 - 1)
<b>SPH</b>	Average level of self-perceived health (range: 1-5, 1 - excellent, 5 - poor)
<b>FI</b>	Average level of food insecurity (0 - no, 1 - yes)
<b>IV</b>	Average level of insufficient vegetable intake (0 - completely sufficient, 1 - insufficient)
<b>SIH</b>	Percentage of smoking inside home (range: 0 - 1)
<b>SBC</b>	Average level of sense of belonging to community (range: 1 - 4, 1 - very strong, 4 - very weak)
<b>SPUH</b>	Percentage of self-perceived unmet health care (range: 0 - 1)
<b>DS</b>	Average daily smokes (Number of cigarettes)
<b>HD</b>	Percentage of heavy drinker (range: 0 - 1)
<b>EU</b>	Average level of emotional unhappiness (range: 0 - 1)
<b>SPS</b>	Average level of self-perceived stress (1 - not stressful, 5 - extremely stressful)
<b>PI</b>	Average level of physical inactiveness (0 - active, 1 - inactive)



## Appendix B

### SAS models for mediational analysis of LBW and preterm births

#### *PRETERM BIRTHS*

##### **MX models:**

```
proc mixed data=mywork.PRETERMfa01;
class dauid01;
model factor1 = NBVAR/solution;
random intercept / subject=dauid01;
run;
proc mixed data=mywork.PRETERMfa01;
class dauid01;
model factor2 = NBVAR/solution;
random intercept / subject=dauid01;
run;
proc mixed data=mywork.PRETERMfa01;
class dauid01;
model factor3 = NBVAR/solution;
random intercept / subject=dauid01;
run;
proc mixed data=mywork.PRETERMfa01;
class dauid01;
model factor5 = NBVAR/solution;
random intercept / subject=dauid01;
run;
```

##### **MX Y model:**

```
proc glimmix ABSPCONV=1e-4 data=mywork.PRETERMfa01;
class dauid01;
model PRETERM (event='1') = age19 age36 female multibirth factor1 factor2 factor3 factor5
NBVAR factor1*NBVAR factor2*NBVAR factor3*NBVAR factor5*NBVAR
/ dist=binary solution;
random intercept factor3 factor5 / subject=dauid01;
id age19 age36 female multibirth factor1 factor2 factor3 factor5 NBVAR dauid01;
output data=mywork.preterm01_NB_xmy pred=p;
run;
```

##### **XY model:**

```
proc mixed data=mywork.preterm01_NB_xmy;
class dauid01;
model p = age19 age36 female multibirth NBVAR/solution;
random intercept / subject=dauid01;
run;
```

## ***LOW BIRTH WEIGHT***

### **MX models:**

```
proc mixed data=mywork.lbwfa01;  
class dauid01;  
model factor1 = NBVAR /solution;  
random intercept / subject=dauid01;  
run;  
proc mixed data=mywork.lbwfa01;  
class dauid01;  
model factor2 = NBVAR /solution;  
random intercept / subject=dauid01;  
run;  
proc mixed data=mywork.lbwfa01;  
class dauid01;  
model factor3 = NBVAR /solution;  
random intercept / subject=dauid01;  
run;  
proc mixed data=mywork.lbwfa01;  
class dauid01;  
model factor4 = NBVAR /solution;  
random intercept / subject=dauid01;  
run;
```

### **MX Y model:**

```
proc glimmix ABSPCONV=1e-4 data=mywork.lbwfa01;  
class dauid01;  
model lbw (event='1') = age19 age36 female multibirth PRETERM factor1 factor2 factor3  
factor4 NBVAR factor1*NBVAR factor2*NBVAR factor3*NBVAR factor4*NBVAR /  
dist=binary solution;  
random intercept factor2 factor4 / subject=dauid01;  
id age19 age36 female multibirth PRETERM factor1 factor2 factor3 factor4 NBVAR  
dauid01;  
output data=mywork.lbw01_NBVAR_xmy pred=p;  
run;
```

### **XY model:**

```
proc mixed data=mywork.lbw01_NBVAR_xmy;  
class dauid01;  
model p = age19 age36 female multibirth PRETERM NBVAR /solution;  
random intercept / subject=dauid01;  
run;
```

**\*where NBVAR is one of the neighbourhood variables**

## Appendix C

### Mediational analysis results for preterm births

<i>NBVAR</i>	<i>Mediator</i>	<i>a</i>	<i>b</i>	<i>sd(a)</i>	<i>sd(b)</i>	<i>ab</i>	<i>sd(ab)</i>
<b>Low Socio-economic status (LOW_SES)</b>	Factor1: SES	1.2917	0.0887	0.0557	0.0166	0.1145	0.0221
	Factor2: Psycho-social	0.3003	0.0706	0.0365	0.0142	0.0212	0.005
	Factor3: Behavioural	0.4284	0.1263	0.0408	0.0179	0.0541	0.0093
	Factor4: Health	0.1429	0.4384	0.0321	0.0171	0.0626	0.0143
	Indirect effect: sum(ab)					0.2525	0.0284
	Direct effect: c'					0.641	0.1838
	Indirect + direct: sum(ab)+c'					0.8935	
						0.882	
<b>Low education (LOW_EDU)</b>	Factor1: SES	0.3284	0.0732	0.0289	0.0172	0.0240	0.0060
	Factor2: Psycho-social	0.1242	0.0638	0.0171	0.0152	0.0079	0.0022
	Factor3: Behavioural	0.1972	0.0976	0.0188	0.0199	0.0193	0.0043
	Indirect effect: sum(ab)					0.0512	0.0078
	Direct effect: c'					0.3052	0.0866
	Indirect + direct: sum(ab)+c'					0.3564	
	Total effect: c					0.3391	
<b>LOW_INCOME</b>	Factor1: SES	1.2374	0.0890	0.0430	0.0167	0.1102	0.0211
	Factor2: Psycho-social	0.1754	0.0713	0.0307	0.0141	0.0125	0.0033
	Factor3: Behavioural	0.1792	0.1275	0.0349	0.0178	0.0228	0.0055
	Factor4: Health	0.1539	0.4391	0.0264	0.0171	0.0676	0.0119
	Indirect effect: sum(ab)					0.2131	0.0250
	Direct effect: c'					0.4249	0.1494
	Indirect + direct: sum(ab)+c'					0.6380	
	Total effect: c					0.6358	
<b>Unemployment rate (UNEMPLOY)</b>	Factor1: SES	0.02021	0.0928	0.0012	0.0166	0.0019	0.0004
	Factor2: Psycho-social	0.0029	0.0729	0.0008	0.0141	0.0002	0.0001
	Factor3: Behavioural	0.0040	0.1284	0.0009	0.0179	0.0005	0.0001
	Factor4: Health	0.0035	0.4402	0.0007	0.0171	0.0015	0.0003
	Indirect effect: sum(ab)					0.0041	0.0005
	Direct effect: c'					0.0073	0.0037
	Indirect + direct: sum(ab)+c'					0.0115	
	Total effect: c					0.0111	
<b>Visible Minority (VIS_MINO)</b>	Factor1: SES	0.877	0.0607	0.0501	0.0161	0.0533	0.0145
	Factor4: Health	0.1829	0.4731	0.0262	0.0238	0.0865	0.0131
	Indirect effect: sum(ab)					0.1398	0.0196

	Direct effect: c'					0.2787	
	Indirect + direct: sum(ab)+c'					0.4185	0.1518
	Total effect: c					0.4084	
	Moderation effect: Factor4*VIS_MINO	-0.3176		0.1397			
<b>IMMIGRANTS</b>	Factor1: SES	0.6407	0.0741	0.0562	0.0164	0.0475	0.0113
	Factor3: Behavioural	-0.1291	0.1149	0.0376	0.0185	-0.0148	0.005
	Factor4: Health	0.1722	0.5052	0.0281	0.0381	0.087	0.0157
	Indirect effect: sum(ab)					0.1197	0.0199
	Direct effect: c'					0.3474	0.1638
	Indirect + direct: sum(ab)+c'					0.467	
	Total effect: c					0.4512	
	Moderation effect: Factor4 *IMMIGRANTS	-0.3042		0.1548			
<b>None Canadian citizen (NON_CITIZEN)</b>	Factor1: SES	1.4976	0.0724	0.1062	0.0163	0.1085	0.0257
	Factor2: Psycho-social	0.2251	0.0538	0.0631	0.0141	0.0121	0.0047
	Factor4: Health	0.3753	0.4369	0.053	0.0173	0.164	0.0241
	Indirect effect: sum(ab)					0.2846	0.0355
	Direct effect: c'					0.2655	0.3052
	Indirect + direct: sum(ab)+c'					0.5501	
	Total effect: c					0.5572	
<b>None official language speaking (NONOFFLANG)</b>	Factor1: SES	0.6036	0.0728	0.0451	0.0164	0.0439	0.0104
	Factor3: Behavioural	-0.0734	0.1155	0.0306	0.0185	-0.0085	0.0038
	Factor4: Health	0.188	0.4381	0.0223	0.0172	0.0824	0.0103
	Indirect effect: sum(ab)					0.1178	0.0151
	Direct effect: c'					0.2159	0.1304
	Indirect + direct: sum(ab)+c'					0.3337	
	Total effect: c					0.327	0.0251
<b>Aboriginal population (ABORIGINAL)</b>	Factor1: SES	1.8855	0.0975	0.2123	0.0164	0.1839	0.0372
	Factor3: Behavioural	0.6060	0.0729	0.1266	0.0139	0.0442	0.0125
	Factor2: Psycho-social	0.9644	0.1280	0.1386	0.0163	0.1234	0.0237
	Factor4: Health	0.2612	0.4405	0.1074	0.0171	0.1151	0.0475
	Indirect effect: sum(ab)					0.4666	0.0660
	Direct effect: c'					0.7275	0.6171
	Indirect + direct: sum(ab)+c'					1.1941	
	Total effect: c					1.1401	0.1130
<b>Average rent rate (RENT_RATE)</b>	Factor1: SES	0.5453	0.0914	0.0195	0.0167	0.0499	0.0093
	Factor2: Psycho-social	0.0755	0.0718	0.014	0.0141	0.0054	0.0015
	Factor3: Behavioural	0.1023	0.1253	0.0156	0.0178	0.0128	0.0027

	Factor4: Health	0.0436	0.4863	0.0121	0.0236	0.0212	0.006
	Indirect effect: sum(ab)					0.0893	0.0114
	Direct effect: c'					0.2284	0.0695
	Indirect + direct: sum(ab)+c'					0.3177	
	Total effect: c					0.313	
	Moderation effect: Factor4 *RENT_RATE	-0.1829		0.0651			
<b>Major house repair (MAJHOSREPA)</b>	Factor1: SES	1.0017	0.0733	0.1045	0.0172	0.0734	0.0189
	Factor2: Psycho-social	0.2093	0.0639	0.0629	0.0152	0.0134	0.0052
	Factor3: Behavioural	0.5629	0.0552	0.0683	0.0307	0.0310	0.0178
	Indirect effect: sum(ab)					0.1178	0.0265
	Direct effect: c'					0.9118	0.3053
	Indirect + direct: sum(ab)+c'					1.0296	
	Total effect: c					1.038	
	Moderation effect: Factor3 *majhosrepa	0.5486		0.2914			
<b>Dwelling value (DWL_VAL)</b>	Factor1: SES	-0.0756	0.0906	0.0062	0.0166	-0.0069	0.0014
	Factor2: Psycho-social	-0.0264	0.0711	0.0039	0.0142	-0.0019	0.0005
	Factor3: Behavioural	-0.0278	0.1273	0.0043	0.0179	-0.0035	0.0007
	Factor4: Health	-0.0182	0.4374	0.0033	0.0171	-0.008	0.0015
	Indirect effect: sum(ab)					-0.0203	0.0022
	Direct effect: c'					-0.0925	0.0203
	Indirect + direct: sum(ab)+c'					-0.1127	
	Total effect: c					-0.1117	
<b>Family income (FAMILY_INC)</b>	Factor1: SES	-0.1022	0.0915	0.0058	0.0164	-0.0094	0.0018
	Factor2: Psycho-social	-0.0276	0.0702	0.0038	0.0139	-0.0019	0.0005
	Factor3: Behavioural	-0.0351	0.1259	0.0042	0.0163	-0.0044	0.0008
	Factor4: Health	-0.0114	0.4380	0.0033	0.0171	-0.0050	0.0015
	Indirect effect: sum(ab)					-0.0207	0.0025
	Direct effect: c'					-0.0781	0.0201
	Indirect + direct: sum(ab)+c'					-0.0987	
	Total effect: c					-0.0973	0.0030
<b>Food insecurity (FI)</b>	Factor2: Psycho-social	0.2463	0.0712	0.0554	0.0139	0.0175	0.0052
	Factor3: Behavioural	0.3835	0.1268	0.0615	0.0163	0.0486	0.0100
	Factor4: Health	0.1973	0.4389	0.0471	0.0171	0.0866	0.0210
	Indirect effect: sum(ab)					0.1527	0.0238
	Direct effect: c'					0.9074	0.2683
	Indirect + direct: sum(ab)+c'					1.0601	
	Total effect: c					1.1963	0.0655

<b>Insufficient vegetable intakes (IV)</b>	Factor1: SES	0.6396	0.0961	0.0773	0.0164	0.0615	0.0128
	Factor2: Psycho-social	0.1129	0.0726	0.0462	0.0139	0.0082	0.0037
	Factor3: Behavioural	0.1883	0.1277	0.0512	0.0163	0.0240	0.0072
	Factor4: Health	0.2274	0.4393	0.0388	0.0171	0.0999	0.0175
	Indirect effect: sum(ab)					0.1936	0.0232
	Direct effect: c'					0.6339	0.2265
	Indirect + direct: sum(ab)+c'					0.8275	
	Total effect: c					0.8324	0.0398
<b>Hard Drinking (HD)</b>	Factor1: SES	-0.4876	0.0979	0.0902	0.0163	-0.0477	0.0119
	Factor2: Psycho-social	-0.1403	0.0731	0.0538	0.0139	-0.0103	0.0044
	Factor3: Behavioural	-0.1835	0.1283	0.0594	0.0163	-0.0235	0.0082
	Factor4: Health	-0.1210	0.4406	0.0453	0.0171	-0.0533	0.0201
	Indirect effect: sum(ab)					-0.1348	0.0251
	Direct effect: c'					-0.4214	0.2629
	Indirect + direct: sum(ab)+c'					-0.5562	
	Total effect: c					-0.5512	0.0466
<b>Chronic Health conditions (CHC)</b>	Factor1: SES	-0.2808	0.0981	0.0780	0.0163	-0.0275	0.0089
	Factor2: Psycho-social	-0.0953	0.0730	0.0448	0.0139	-0.0070	0.0035
	Factor4: Health	-0.2639	0.4389	0.0372	0.0171	-0.1158	0.0169
	Indirect effect: sum(ab)					-0.1503	0.0195
	Direct effect: c'					-0.6402	0.2176
	Indirect + direct: sum(ab)+c'					-0.7905	
	Total effect: c					-0.7839	0.0398
<b>Self-perceived health (SPH)</b>	Factor2: Psycho-social	-0.0255	0.0721	0.0092	0.0140	-0.0018	0.0008
	Factor4: Health	-0.0647	0.7209	0.0077	0.1029	-0.0466	0.0086
	Indirect effect: sum(ab)					-0.0485	0.0087
	Direct effect: c'					-0.0427	0.0452
	Indirect + direct: sum(ab)+c'					-0.0912	
	Total effect: c					-0.0747	0.0087
	Moderation effect: Factor4 *sph	-0.1230		0.0446			

## Appendix D

### Mediational analysis results for LBW

<i>NBVAR</i>	<i>Mediator</i>	<i>a</i>	<i>b</i>	<i>sd(a)</i>	<i>sd(b)</i>	<i>ab</i>	<i>sd(ab)</i>
<b>Not a Canadian citizen (NO_CITIZEN)</b>	Factor1: Behavioural	0.3947	0.16	0.0891	0.0206	0.0632	0.0165
	Factor2: SES	1.3509	0.0774	0.0975	0.0240	0.1045	0.0334
	Factor3: Psycho-social	0.3065	0.1272	0.0622	0.0171	0.039	0.0095
	Factor4: Health	0.2211	0.2414	0.0526	0.0227	0.0534	0.0137
	Indirect effect: sum(ab)					0.2600	0.0408
	Direct effect: c'					1.1011	0.3723
	Indirect + direct: sum(ab)+c'					1.3611	
	Total effect: c					1.2945	0.0604
<b>IMMIGRANTS</b>	Factor4: Health	0.1028	0.2414	0.0277	0.0233	0.0248	0.0071
	Indirect effect: sum(ab)					0.0248	0.0071
	Direct effect: c'					0.6533	0.2036
	Indirect + direct: sum(ab)+c'					0.6781	
	Total effect: c					0.69	0.0312
<b>Visible minority (VIS_MINO)</b>	Factor2: SES	0.8756	0.0385	0.0445	0.0252	0.0337	0.0222
	Factor3: Psycho-social	0.1024	0.0942	0.0308	0.0176	0.0097	0.0035
	Factor4: Health	0.1706	0.2449	0.0255	0.0232	0.0418	0.0074
	Indirect effect: sum(ab)					0.0851	0.0237
	Direct effect: c'					0.7302	0.1856
	Indirect + direct: sum(ab)+c'					0.8153	
	Total effect: c					0.8031	0.0291
<b>ABORIGINAL</b>	Factor1: Behavioural	1.2183	0.1617	0.1694	0.0206	0.197	0.0373
	Factor2: SES	1.361	0.0807	0.197	0.0239	0.1098	0.0365
	Factor3: Psycho-social	0.7438	0.1287	0.1239	0.017	0.0957	0.0205
	Factor4: Health	0.3599	0.2424	0.1043	0.0227	0.0872	0.0267
	Indirect effect: sum(ab)					0.4898	0.0621
	Direct effect: c'					1.1381	0.7741
	Indirect + direct: sum(ab)+c'					1.6279	
	Total effect: c					1.5623	0.1193
<b>Official language not English (NONOFFLANG)</b>	Factor1: Behavioural	0.1338	0.1603	0.0377	0.0206	0.0214	0.0067
	Factor2: SES	0.5264	0.0767	0.0417	0.024	0.0404	0.0130
	Factor3: Psycho-social	0.105	0.1276	0.0266	0.017	0.0134	0.0039
	Factor4: Health	0.1096	0.2412	0.0223	0.0227	0.0264	0.006
	Indirect effect: sum(ab)					0.1017	0.0163
	Direct effect: c'					0.5626	0.1613

	Indirect + direct: sum(ab)+c'					0.6643	
	Total effect: c					0.6424	0.025
<b>Dwelling value (DWL_VAL)</b>	Factor1: Behavioural	-0.0526	0.1886	0.005	0.024	-0.0099	0.0016
	Factor2: SES	-0.0522	0.0786	0.0059	0.0238	-0.0041	0.0013
	Factor3: Psycho-social	-0.0384	0.1288	0.0038	0.0172	-0.0049	0.0008
	Factor4: Health	-0.0158	0.2393	0.0033	0.0227	-0.0038	0.0009
	Indirect effect: sum(ab)					-0.0227	0.0024
	Direct effect: c'					-0.0717	0.026
	Indirect + direct: sum(ab)+c'					-0.0944	
	Total effect: c					-0.0916	0.0034
	Moderation effect: Factor1 *DWL_VAL	0.0604		0.027			
<b>Percentage of movers (MOVERS)</b>	Factor1: Behavioural	0.1328	0.1623	0.0285	0.0213	0.0216	0.0055
	Factor2: SES	0.4	0.0625	0.0318	0.0244	0.0250	0.01
	Factor3: Psycho-social	0.0503	0.1172	0.0198	0.0172	0.0059	0.0025
	Indirect effect: sum(ab)					0.0525	0.0116
	Direct effect: c'					0.3615	0.1195
	Indirect + direct: sum(ab)+c'					0.414	
	Total effect: c					0.4071	0.0209
<b>Average Person per room (PERROOM)</b>	Factor1: Behavioural	-0.0823	0.1602	0.0162	0.0213	-0.0132	0.0032
	Factor2: SES	-0.2207	0.0606	0.0179	0.0244	-0.0134	0.0055
	Factor3: Psycho-social	-0.0368	0.1162	0.0119	0.0173	-0.0043	0.0015
	Indirect effect: sum(ab)					-0.0309	0.0065
	Direct effect: c'					-0.3025	0.0736
	Indirect + direct: sum(ab)+c'					-0.3334	
	Total effect: c					-0.323	0.01
<b>Average person per bedroom (PERBEDROOM)</b>	Factor1: Behavioural	-0.3057	0.1585	0.0432	0.0206	-0.0485	0.0093
	Factor2: SES	-0.6181	0.0764	0.048	0.0239	-0.0472	0.0153
	Factor3: Psycho-social	-0.1548	0.1265	0.0319	0.0171	-0.0196	0.0049
	Factor4: Health	-0.0749	0.2407	0.0268	0.0228	-0.018	0.0067
	Indirect effect: sum(ab)					-0.1333	0.0197
	Direct effect: c'					-0.8446	0.1984
	Indirect + direct: sum(ab)+c'					-0.9779	
	Total effect: c					-0.9489	0.0273
<b>RENT_RATE</b>	Factor1: Behavioural	0.1808	0.1588	0.0188	0.0207	0.0287	0.0048
	Factor2: SES	0.4741	0.0719	0.0184	0.024	0.0341	0.0114
	Factor3: Psycho-social	0.0974	0.1227	0.0137	0.0174	0.012	0.0024
	Factor4: Health	0.0246	0.3008	0.0118	0.0309	0.0073	0.0036



	Indirect effect: sum(ab)					0.0822	0.0132
	Direct effect: c'					0.4149	0.0857
	Indirect + direct: sum(ab)+c'					0.497	
	Total effect: c					0.4803	0.0118
	Moderation effect: Factor4 *RENT_RATE	-0.2273		0.0829			
<b>Total income (TOT_INC)</b>	Factor1: Behavioural	-0.0608	0.1592	0.0049	0.0214	-0.0097	0.0015
	Factor2: SES	-0.0705	0.0611	0.0057	0.0243	-0.004	0.0018
	Factor3: Psycho-social	-0.0356	0.1333	0.0037	0.0206	-0.0047	0.0009
	Indirect effect: sum(ab)					-0.0187	0.0025
	Direct effect: c'					-0.0934	0.0256
	Indirect + direct: sum(ab)+c'					-0.112	
	Total effect: c					-0.1068	0.0032
<b>Family income (FAMILY_INC)</b>	Factor1: Behavioural	-0.061	0.157	0.0049	0.0207	-0.0096	0.0015
	Factor2: SES	-0.0782	0.0756	0.0056	0.0239	-0.0059	0.0019
	Factor3: Psycho-social	-0.0375	0.1252	0.0037	0.0171	-0.0047	0.0008
	Factor4: Health	-0.0078	0.2404	0.0032	0.0227	-0.0019	0.0008
	Indirect effect: sum(ab)					-0.0221	0.0027
	Direct effect: c'					-0.0919	0.0259
	Indirect + direct: sum(ab)+c'					-0.114	
	Total effect: c					-0.109	0.0033
<b>Percentage of low income (LOW_INCOME)</b>	Factor1: Behavioural	0.3625	0.159	0.0425	0.0207	0.0576	0.0101
	Factor2: SES	1.0831	0.0736	0.0402	0.0241	0.0797	0.0263
	Factor3: Psycho-social	0.2349	0.1243	0.0299	0.0172	0.0292	0.0055
	Factor4: Health	0.1213	0.2886	0.0258	0.0306	0.035	0.0084
	Indirect effect: sum(ab)					0.2015	0.0299
	Direct effect: c'					0.7036	0.1862
	Indirect + direct: sum(ab)+c'					0.9051	
	Total effect: c					0.8488	0.0287
	Moderation effect: Factor4 *LOW_INCOME	-0.381		0.1677			
<b>Unemployment rate (UNEMPLOY)</b>	Factor1: Behavioural	0.0062	0.1598	0.001	0.0206	0.001	0.0002
	Factor2: SES	0.0183	0.077	0.0011	0.0241	0.0014	0.0005
	Factor3: Psycho-social	0.0038	0.1274	0.0008	0.0171	0.0005	0.0001
	Factor4: Health	0.0033	0.2415	0.0006	0.0227	0.0008	0.0002
	Indirect effect: sum(ab)					0.0037	0.0005
	Direct effect: c'					0.0119	0.0047
	Indirect + direct: sum(ab)+c'					0.0156	
	Total effect: c					0.0145	0.0007

<b>Percentage of low education (LOW_EDU)</b>	Factor1: Behavioural	0.3581	0.289	0.0219	0.069	0.1035	0.0255
	Factor2: SES	0.1848	0.0636	0.0275	0.0243	0.0118	0.0049
	Factor3: Psycho-social	0.1721	0.117	0.0166	0.0173	0.0201	0.0036
	Indirect effect: sum(ab)					0.1354	0.0262
	Direct effect: c'					0.2901	0.1096
	Indirect + direct: sum(ab)+c'					0.4255	
	Total effect: c					0.3541	0.0159
<b>Low socio-economic status (LOW_SES)</b>	Factor1: Behavioural	0.7954	0.1555	0.0476	0.0207	0.1237	0.0181
	Factor2: SES	0.9676	0.075	0.0548	0.024	0.0725	0.0236
	Factor3: Psycho-social	0.412	0.1247	0.0352	0.0172	0.0514	0.0083
	Factor4: Health	0.0734	0.2406	0.0316	0.0227	0.0176	0.0078
	Indirect effect: sum(ab)					0.2652	0.0319
	Direct effect: c'					0.8076	0.2324
	Indirect + direct: sum(ab)+c'					1.0728	
	Total effect: c					1.0231	0.034
<b>Emotional unhappiness (EU)</b>	Factor1: Behavioural	1.0845	0.1609	0.2544	0.0206	0.1745	0.0466
	Factor2: SES	1.7896	0.0800	0.2950	0.0239	0.1431	0.0489
	Factor3: Psycho-social	0.5963	0.1287	0.1852	0.0170	0.0767	0.0259
	Indirect effect: sum(ab)					0.3944	0.0724
	Direct effect: c'					3.0583	1.1293
	Indirect + direct: sum(ab)+c'					3.4527	
	Total effect: c					3.4092	0.1707
<b>Sense of belonging to community (SBC)</b>	Factor1: Behavioural	0.0935	0.1614	0.0333	0.0206	0.0151	0.0057
	Factor2: SES	0.3845	0.0794	0.0374	0.0240	0.0305	0.0097
	Factor4: Health	0.0505	0.2429	0.0202	0.0227	0.0123	0.0050
	Indirect effect: sum(ab)					0.0579	0.0123
	Direct effect: c'					0.2551	0.1496
	Indirect + direct: sum(ab)+c'					0.3130	
	Total effect: c					0.2995	0.0232
<b>Food insecurity (FI)</b>	Factor1: Behavioural	0.7895	0.1572	0.0734	0.0207	0.1241	0.0200
	Factor2: SES	1.2810	0.0759	0.0820	0.0240	0.0973	0.0313
	Factor3: Psycho-social	0.3556	0.1262	0.0540	0.0171	0.0449	0.0091
	Factor4: Health	0.1969	0.2406	0.0460	0.0228	0.0474	0.0119
	Indirect effect: sum(ab)					0.3136	0.0401
	Direct effect: c'					1.2004	0.3356
	Indirect + direct: sum(ab)+c'					1.5140	
	Total effect: c					1.4386	0.0504
<b>Insufficient</b>	Factor1: Behavioural	0.3139	0.1606	0.0619	0.0206	0.0504	0.0119

<b>vegetable intake (IV)</b>	Factor2: SES	0.4943	0.0788	0.0715	0.0240	0.0390	0.0131
	Factor3: Psycho-social	0.1718	0.1280	0.0455	0.0171	0.0220	0.0065
	Factor4: Health	0.2517	0.2409	0.0377	0.0227	0.0606	0.0107
	Indirect effect: sum(ab)					0.1720	0.0217
	Direct effect: c'					0.8861	0.2855
	Indirect + direct: sum(ab)+c'					1.0581	
	Total effect: c					1.0329	0.0411
<b>Self-perceived unmet health (SPUH)</b>	Factor2: SES	0.7141	0.0800	0.1197	0.0240	0.0572	0.0196
	Factor3: Psycho-social	0.1566	0.1291	0.0749	0.0170	0.0202	0.0100
	Factor4: Health	0.2068	0.2425	0.0626	0.0227	0.0501	0.0159
	Indirect effect: sum(ab)					0.1275	0.0272
	Direct effect: c'					0.8861	0.4620
	Indirect + direct: sum(ab)+c'					1.0136	
	Total effect: c					0.9991	0.0715
<b>Physical inactiveness (PI)</b>	Factor2: SES	0.4985	0.0798	0.0684	0.0239	0.0398	0.0131
	Indirect effect: sum(ab)					0.0398	0.0131
	Direct effect: c'					0.4712	0.2640
	Indirect + direct: sum(ab)+c'					0.5110	
	Total effect: c					0.5141	0.0415
<b>Heavy Drinking (HD)</b>	Factor1: Behavioural	-0.3293	0.1616	0.0718	0.0206	-0.0532	0.0134
	Factor2: SES	-0.3082	0.0813	0.0836	0.0239	-0.0251	0.0100
	Factor2: SES	-0.1943	0.1288	0.0532	0.0170	-0.0250	0.0076
	Factor4: Health	-0.1260	0.2423	0.0444	0.0227	-0.0305	0.0111
	Indirect effect: sum(ab)					-0.1338	0.0215
	Direct effect: c'					-0.8514	0.3360
	Indirect + direct: sum(ab)+c'					-0.9852	
	Total effect: c					-0.9496	0.0494
<b>Regular Drinking (RD)</b>	Factor1: Behavioural	-0.4080	0.1610	0.0625	0.0206	-0.0657	0.0131
	Factor2: SES	-0.6191	0.0798	0.0716	0.0239	-0.0494	0.0159
	Factor3: Psycho-social	-0.2081	0.1287	0.0463	0.0170	-0.0268	0.0069
	Factor4: Health	-0.0845	0.2426	0.0389	0.0227	-0.0205	0.0096
	Indirect effect: sum(ab)					-0.1624	0.0237
	Direct effect: c'					-0.5518	0.2915
	Indirect + direct: sum(ab)+c'					-0.7142	
	Total effect: c					-0.6760	0.0429

## Appendix E

### WinBUGS Code for the random effect models

#### ***PRETERM BIRTHS***

```

model
{
  for (i in 1:m)
  {
    preterm[i] ~ dbern (p[i])
    logit(p[i]) <- gamma01*age19[i]+gamma02*age36[i]+gamma03*female[i]+
      gamma04* multibirth[i]+ alpha0[area[i]]
  }
  for (j in 1:k)
  {
    alpha0[j] <- gamma00+u0[j]+v0[j]
    v0[j]~dnorm(0,tau.v0)
  }
  gamma01~dnorm(0.0, 0.01)
  gamma02~dnorm(0.0, 0.01)
  gamma03~dnorm(0.0, 0.01)
  gamma04~dnorm(0.0, 0.01)
  gamma00~dflat()

  # CAR prior distribution for spatial correlated heterogeneity
  u0[1:k]~car.normal(adj[],weights[],num[],tau.u0)

  #weights
  for(n in 1:sumnumneigh)
  {
    weights[n]<-1.0
  }

  # prior 1 for variances
  tau.v0<-1/(sigma.v0*sigma.v0)
  tau.u0~dgamma(0.01,0.01)
  sigma.v0~dunif(0,100)
  # prior 2 for variances
  #tau.v0< dgamma(0.5,0.0005)
  #tau.u0~dgamma(0.5,0.0005)

  for (l in 1:k)
  {
    uv0[l]<-u0[l]+v0[l]
  }
  varianceu0 <- 1/tau.u0
  variancev0 <- 1/tau.v0
  varianceuv0 <- sd(uv0[])*sd(uv0[])
}

```

## ***LOW BIRTH WEIGHT***

```
model
{
  for (i in 1:m)
  {
    LBW[i] ~ dbern (p[i])
    logit(p[i]) <- gamma01*age19[i]+gamma02*age36[i]+gamma03*female[i]+
      gamma04* multibirth[i]+gamma05*preterm[i]+alpha0[area[i]]
  }
  for (j in 1:k)
  {
    alpha0[j] <- gamma00+u0[j]+v0[j]
    v0[j]~dnorm(0,tau.v0)
  }

  gamma01~dnorm(0.0, 0.01)
  gamma02~dnorm(0.0, 0.01)
  gamma03~dnorm(0.0, 0.01)
  gamma04~dnorm(0.0, 0.01)
  gamma00~dflat()

  # CAR prior distribution for spatial correlated heterogeneity
  u0[1:k]~car.normal(adj[],weights[],num[],tau.u0)

  #weights
  for(n in 1:sumnumneigh)
  {
    weights[n]<-1.0
  }

  # prior 1 for variances
  tau.v0<-1/(sigma.v0*sigma.v0)
  tau.u0~dgamma(0.01,0.01)
  sigma.v0~dunif(0,100)
  # prior 2 for variances
  #tau.v0< dgamma(0.5,0.0005)
  #tau.u0~dgamma(0.5,0.0005)

  for (l in 1:k)
  {
    uv0[l]<-u0[l]+v0[l]
  }

  varianceu0 <- 1/tau.u0
  variancev0 <- 1/tau.v0
  varianceuv0 <- sd(uv0[])*sd(uv0[])
}
```

## Appendix F

### Model convergence testing for the random effect preterm birth model

#### Node statistics

node	mean	sd	MC error	2.5%	median	97.5%	start	sample
gamma00	-2.778	0.02433	4.81E-04	-2.826	-2.777	-2.73	1000	20002
gamma01	0.2356	0.06501	6.98E-04	0.1063	0.2361	0.3625	1000	20002
gamma02	0.2282	0.04326	3.84E-04	0.1419	0.2286	0.3128	1000	20002
gamma03	-0.1318	0.02875	3.54E-04	-0.1883	-0.1316	-0.07559	1000	20002
gamma04	2.964	0.04284	5.72E-04	2.88	2.964	3.047	1000	20002

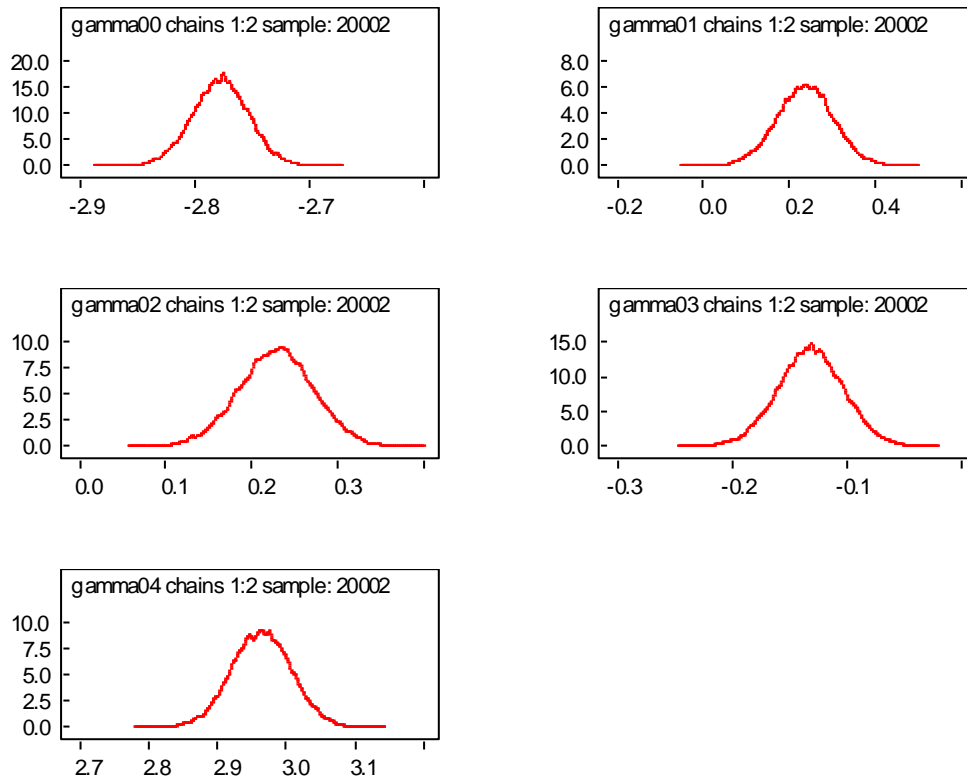
#### DIC – prior 1

	Dbar	Dhat	pD	DIC
preterm	38882.3	38592.7	289.52	39171.8
total	38882.3	38592.7	289.52	39171.8

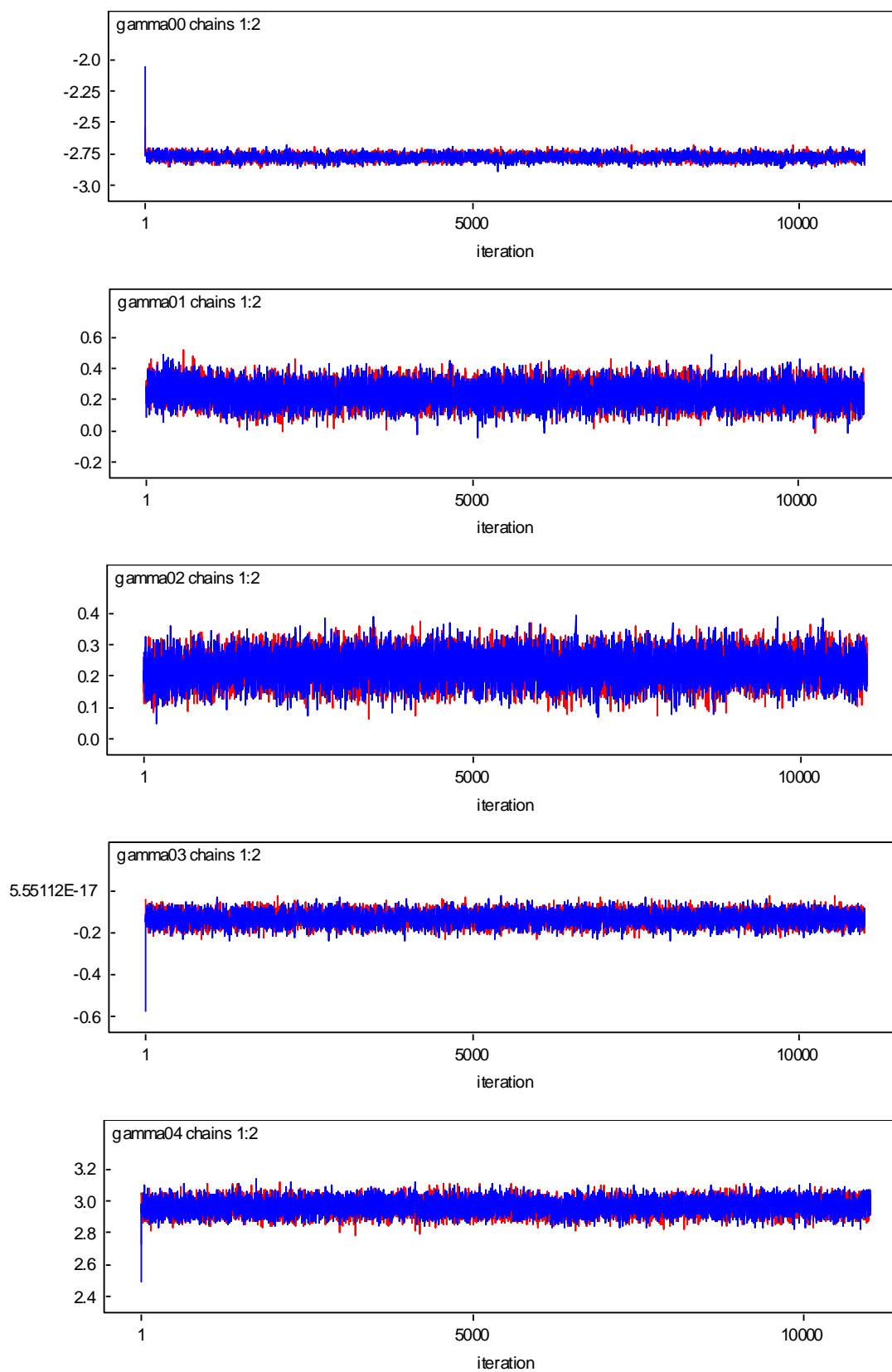
#### DIC – prior 2

	Dbar	Dhat	pD	DIC
Preterm	38898.6	38707.8	190.837	39089.5
Total	38898.6	38707.8	190.837	39089.5

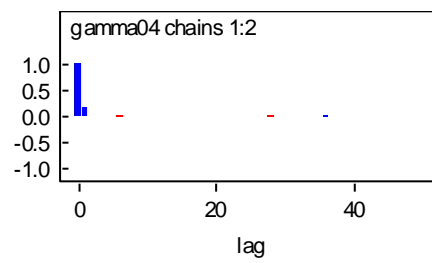
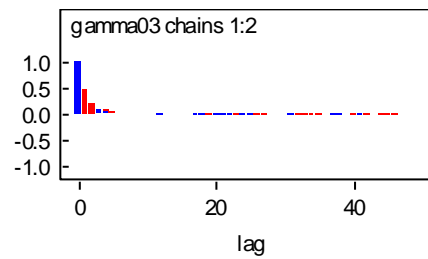
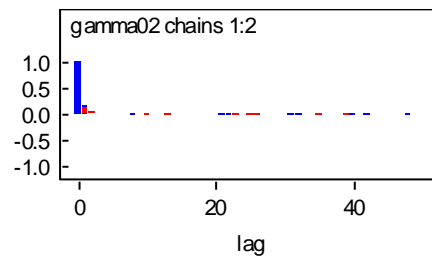
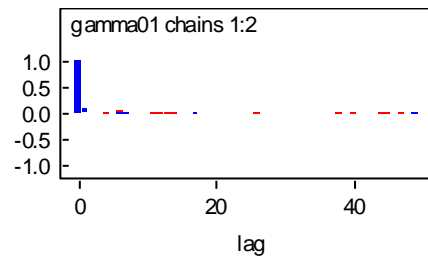
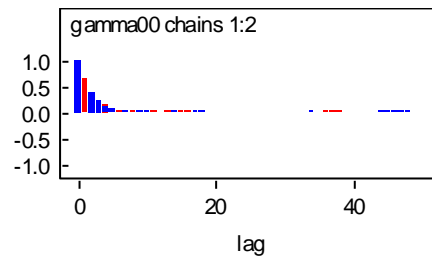
#### Density



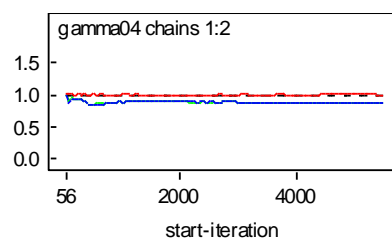
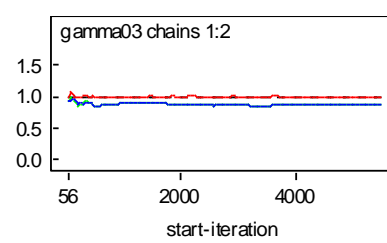
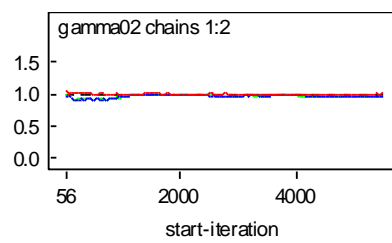
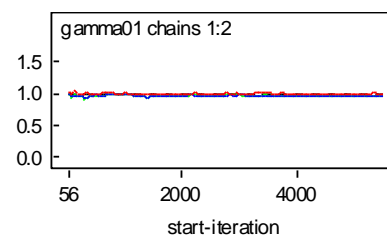
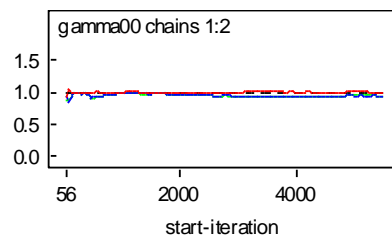
## History



## Autocorrelation



## Gelman Rubin statistic





## Appendix G

### Spatial distribution of personal risk factors of adverse birth outcomes

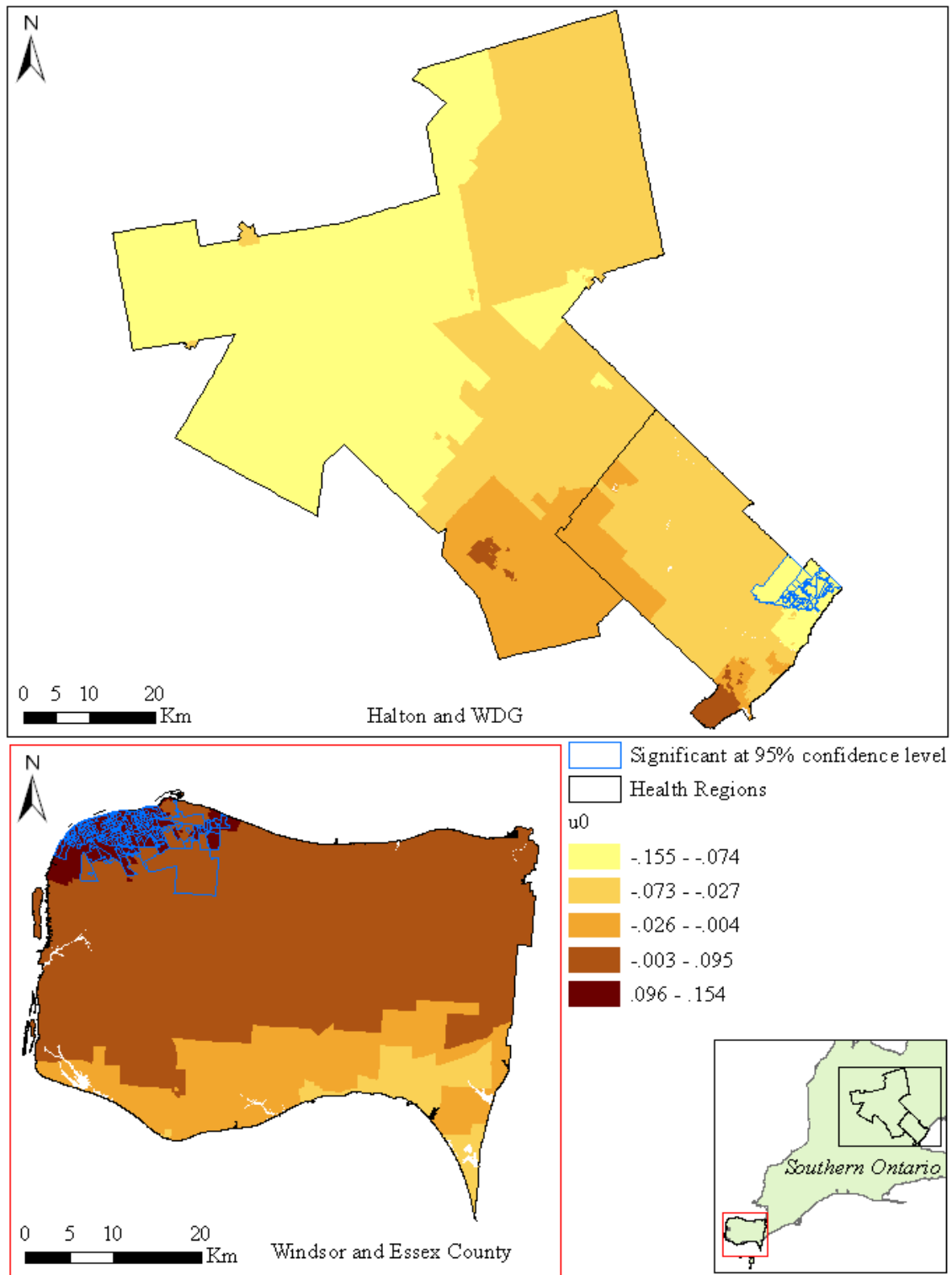


Figure G.1: Spatially correlated random effect  $u_0$  for preterm birth

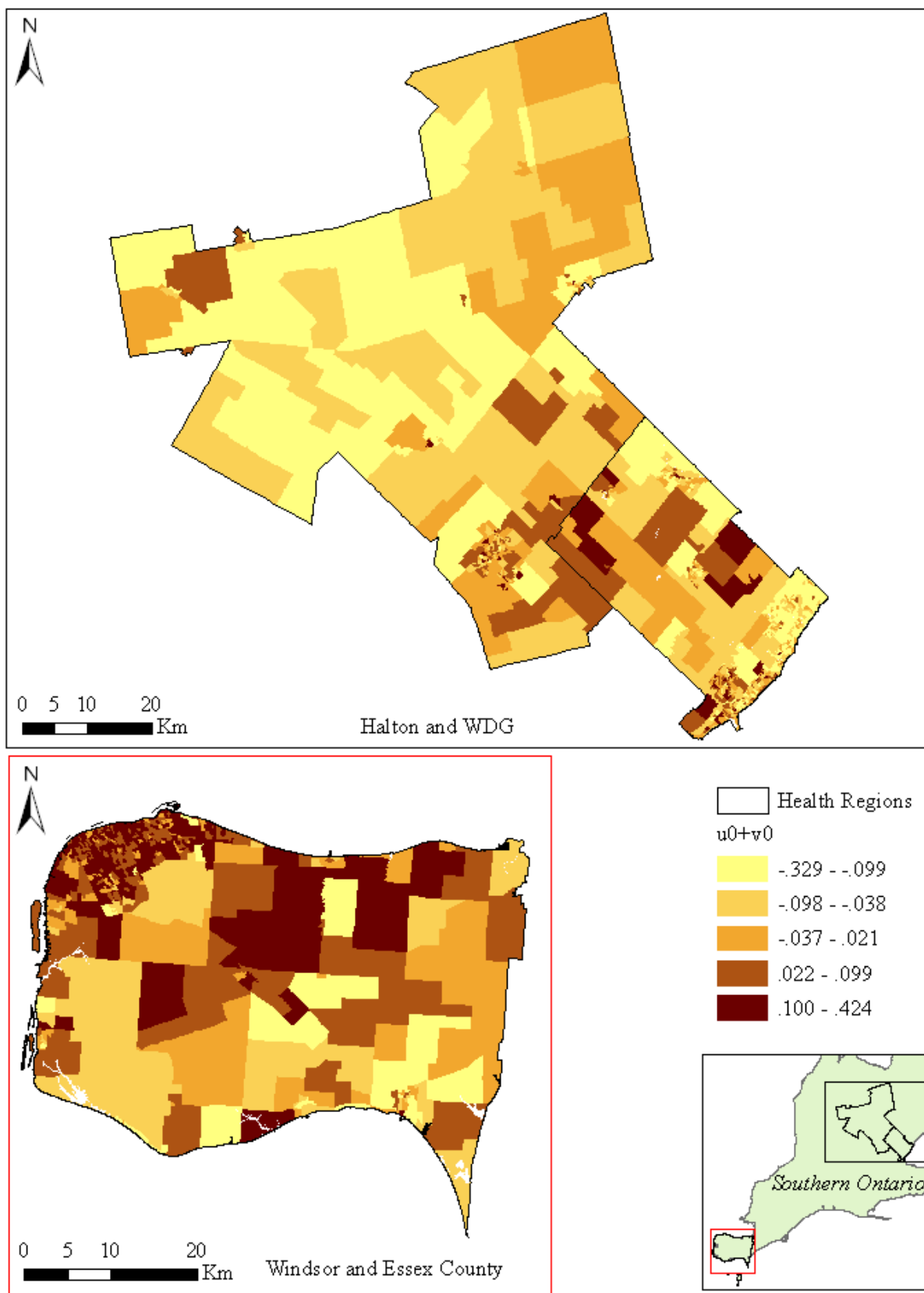


Figure G.2: Total random effects  $u_0 + v_0$  for preterm birth

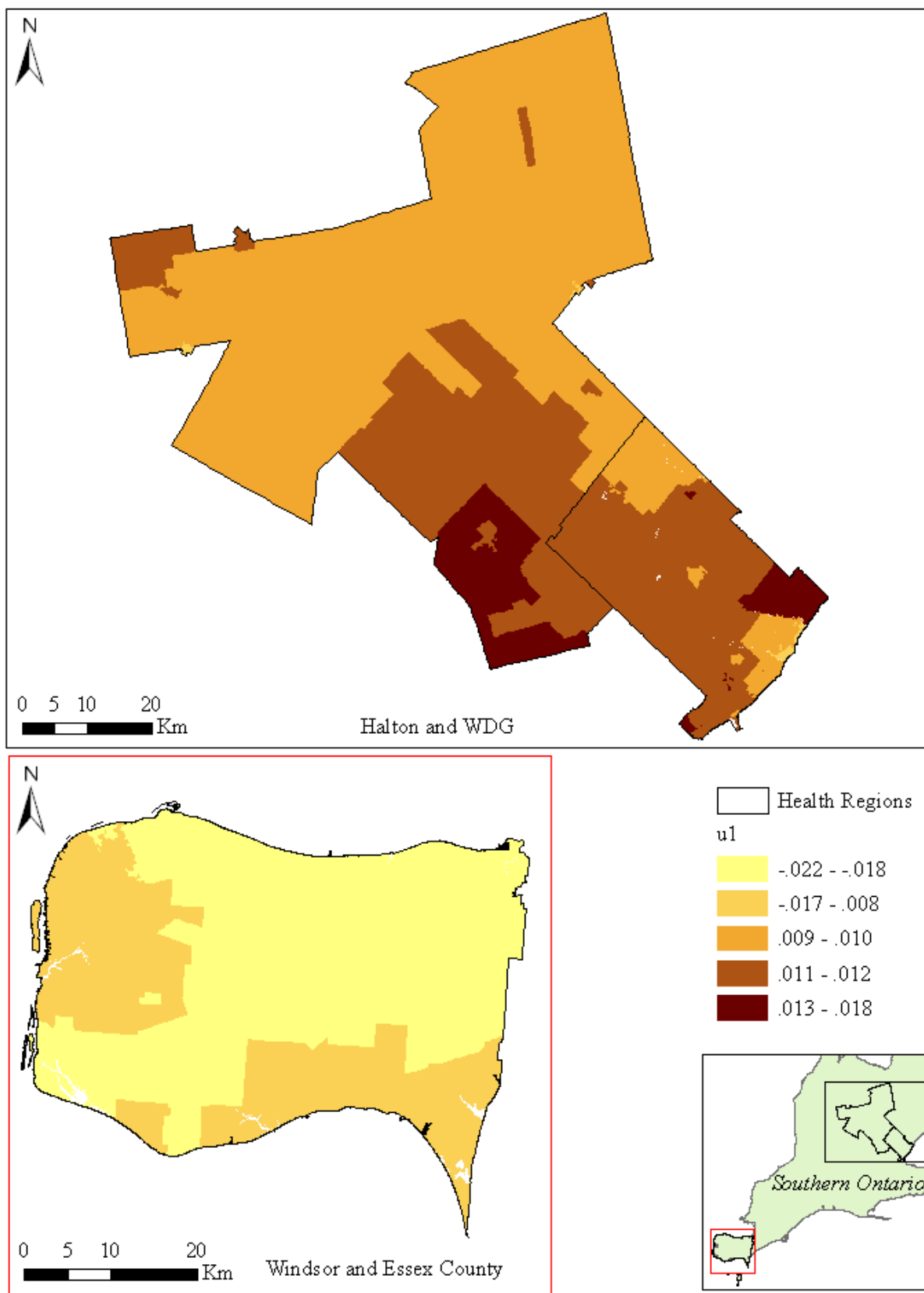


Figure G.3: Spatially correlated random effect of Factor 1 (SES)  $u_l$  for preterm birth

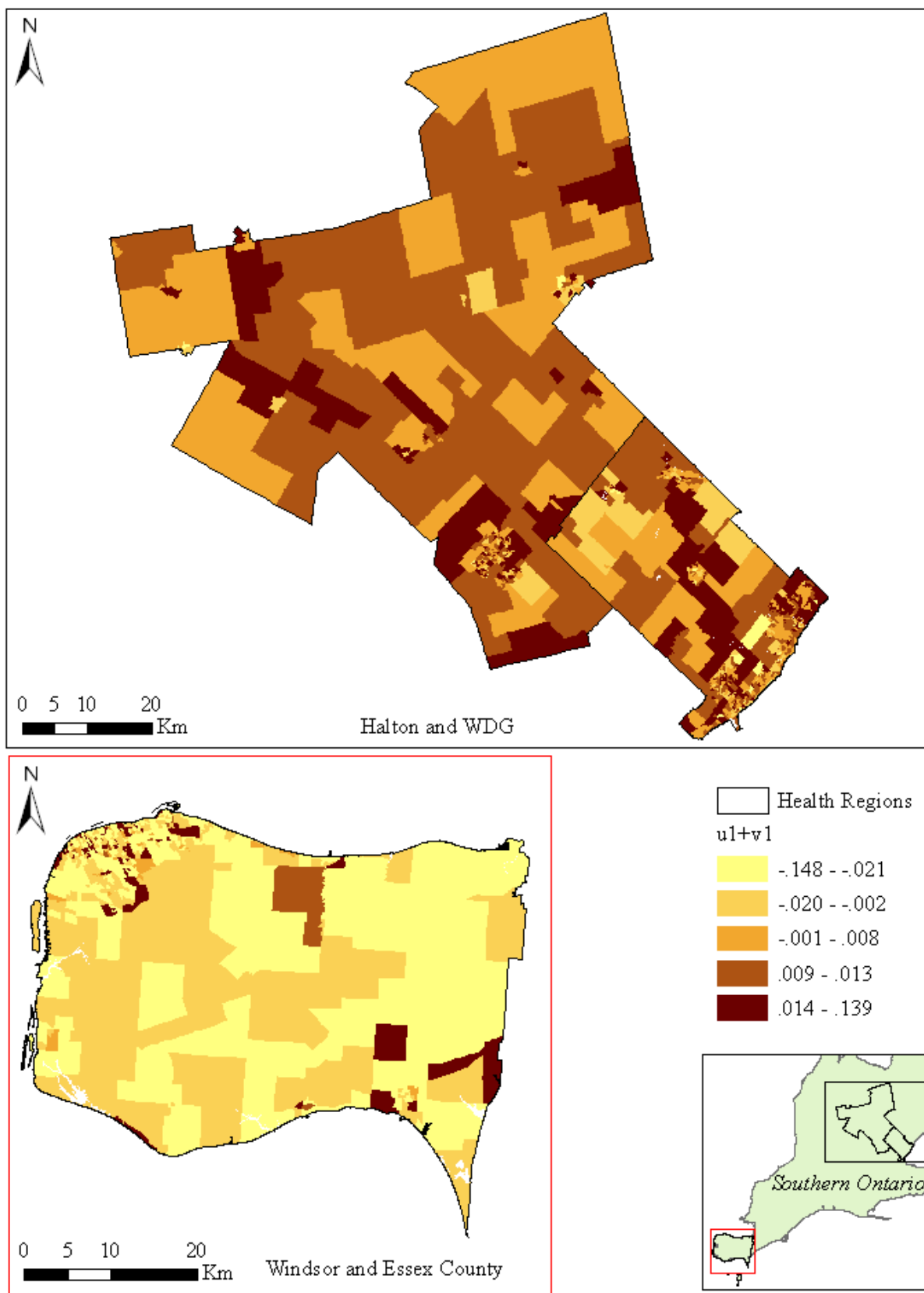


Figure G.4: Total random effects of Factor 1 (SES)  $u_l + v_l$  for preterm birth

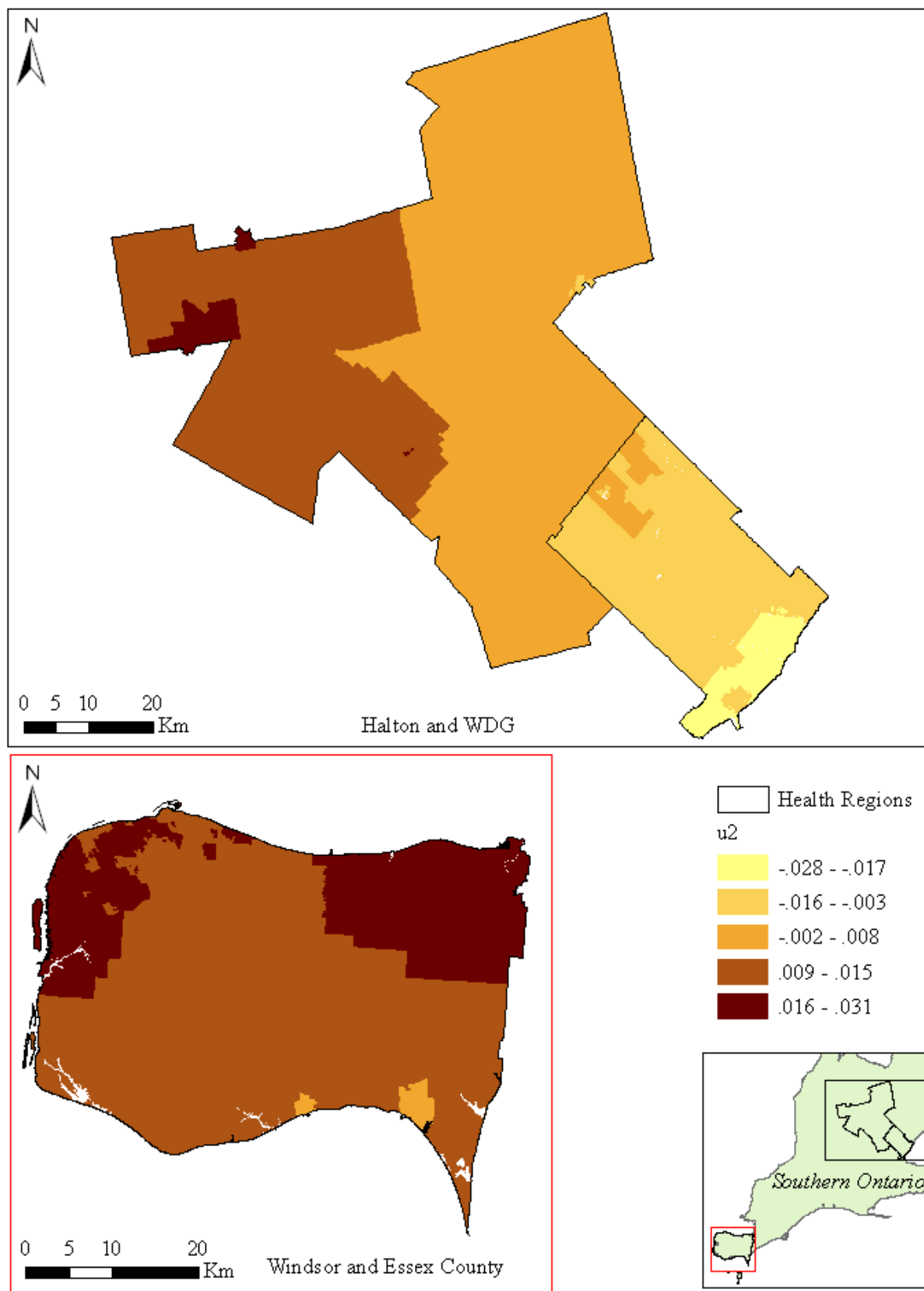


Figure G.5: Spatially correlated random effect of Factor 2 (psycho-social)  $u_2$  for preterm birth

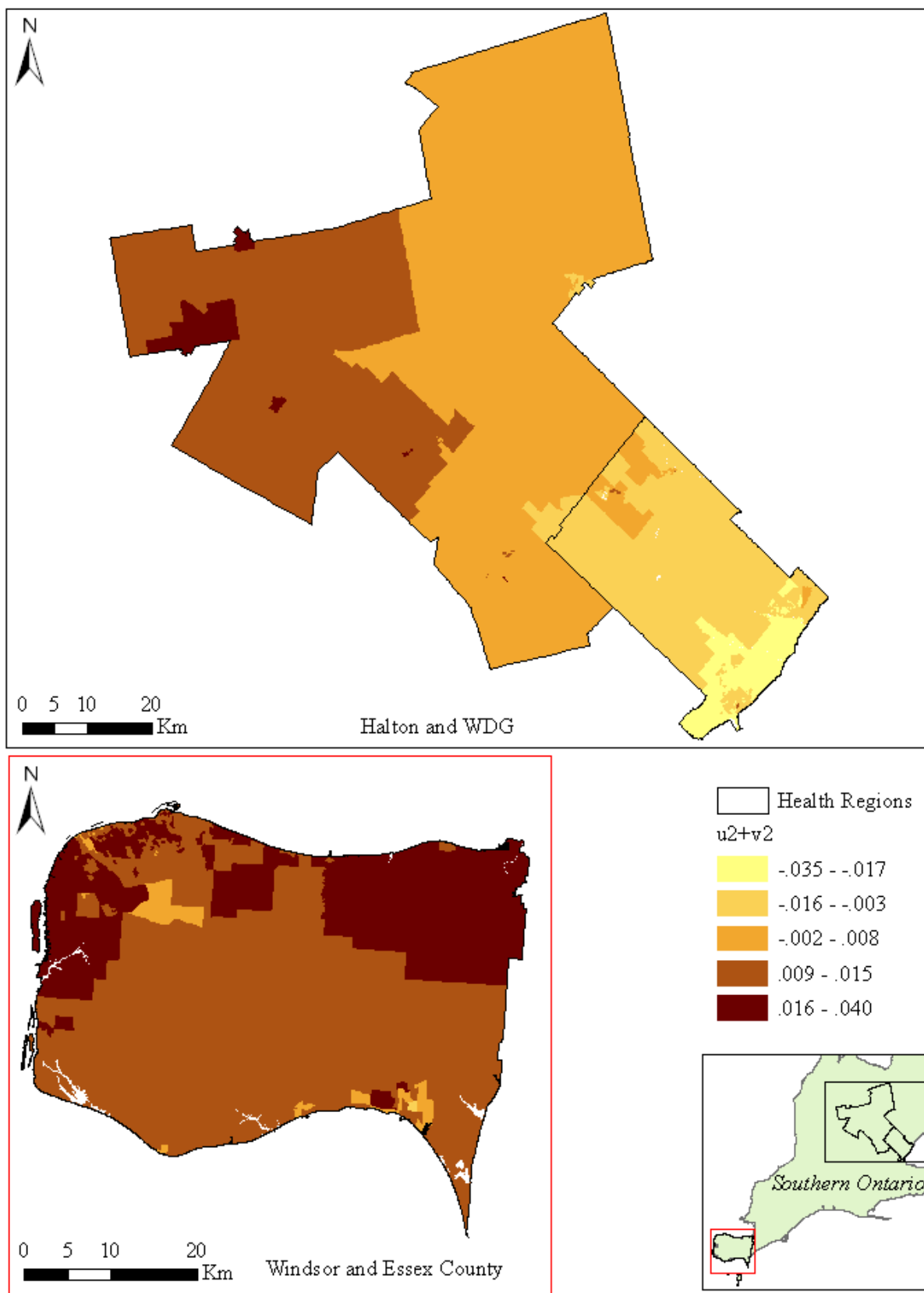


Figure G.6: Total random effects of Factor 2 (psycho-social)  $u_2+v_2$  for preterm birth

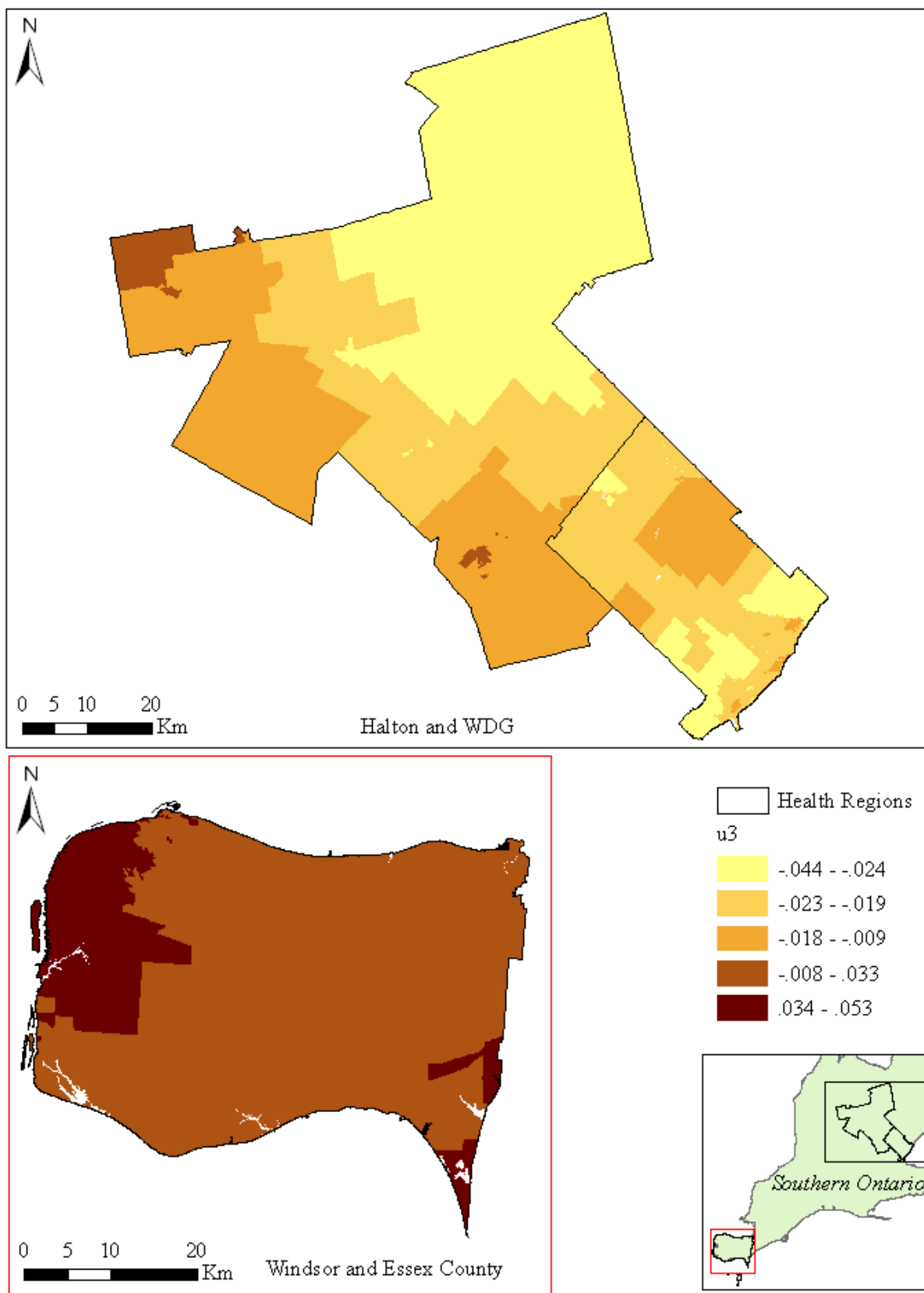


Figure G.7: Spatially correlated random effect of Factor 3 (behavioural)  $u_3$  for preterm birth

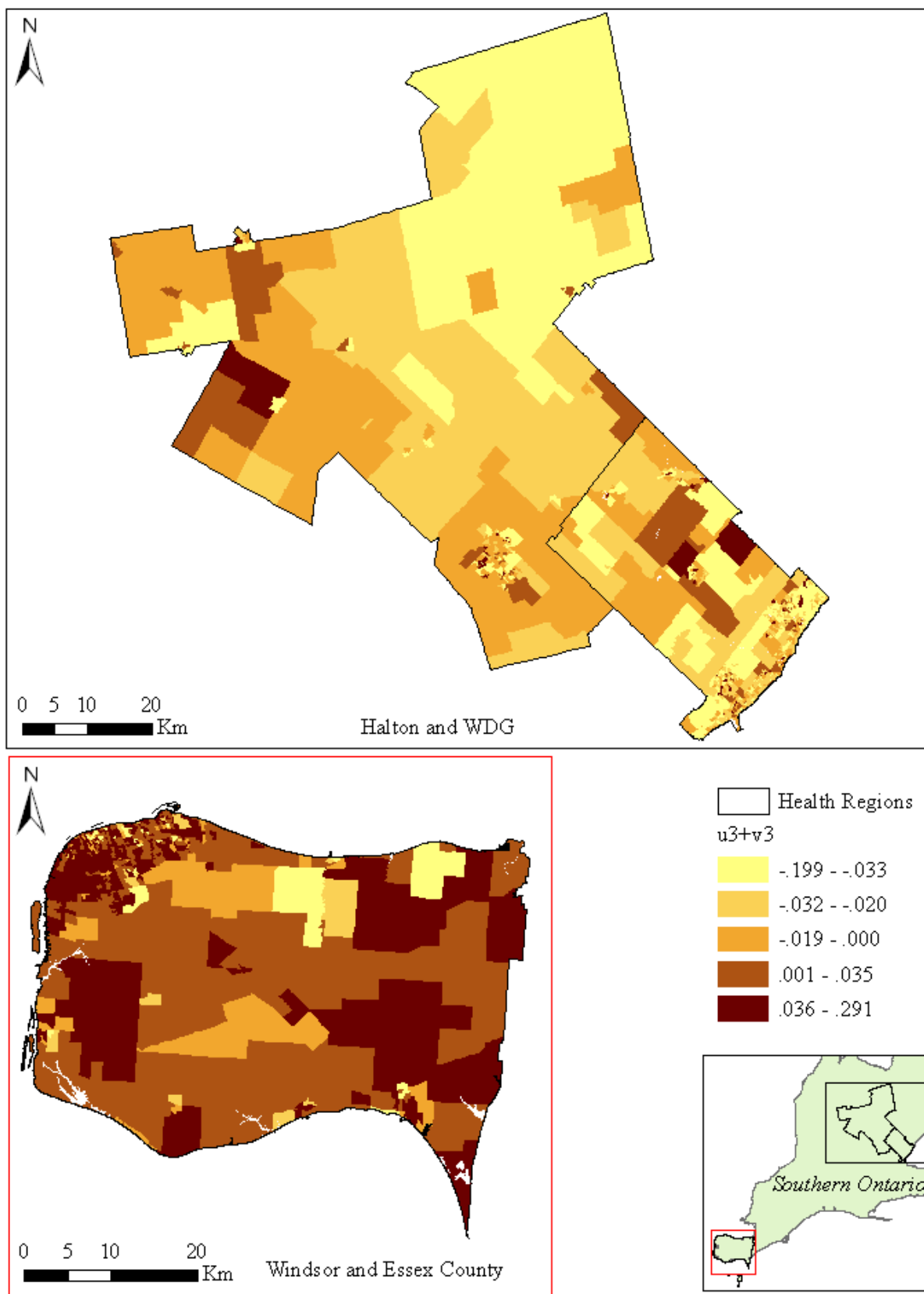


Figure G.8: Total random effects of Factor 3 (behavioural)  $u_3+v_3$  for preterm birth



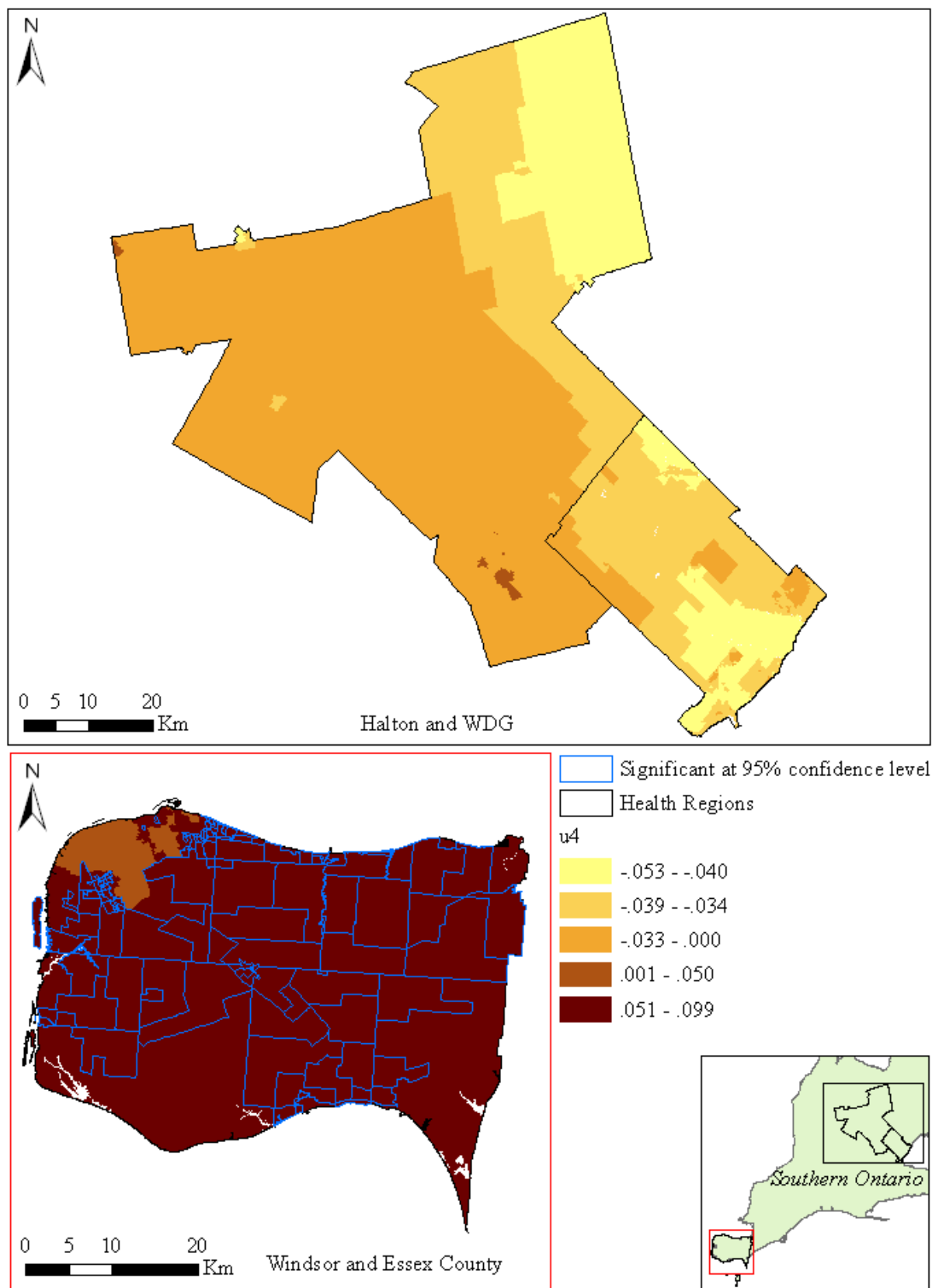


Figure G.9: Spatially correlated random effect of Factor 4 (health)  $u_4$  for preterm birth

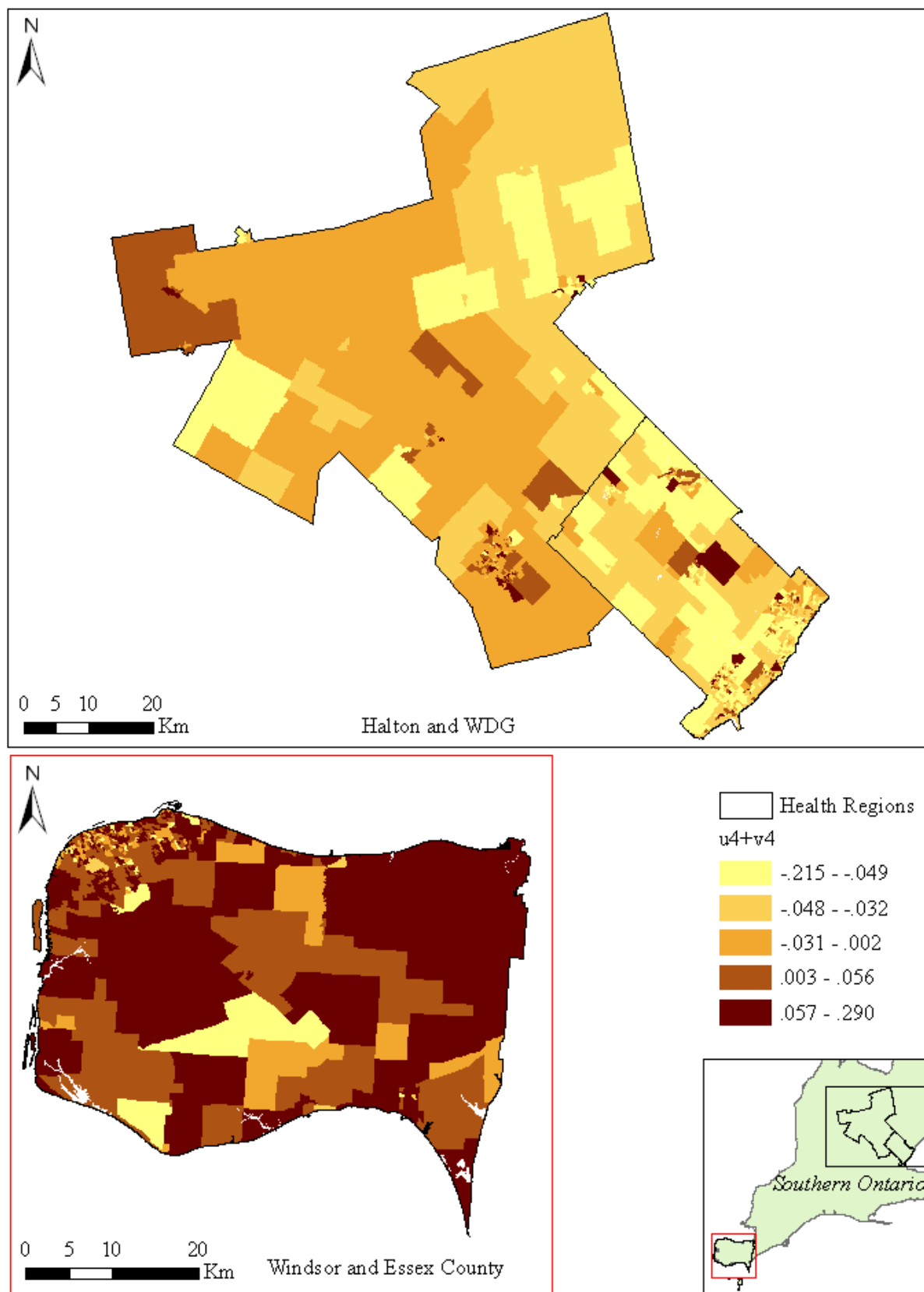


Figure G.10: Total random effects of Factor 4 (health)  $u_4 + v_4$  for preterm birth

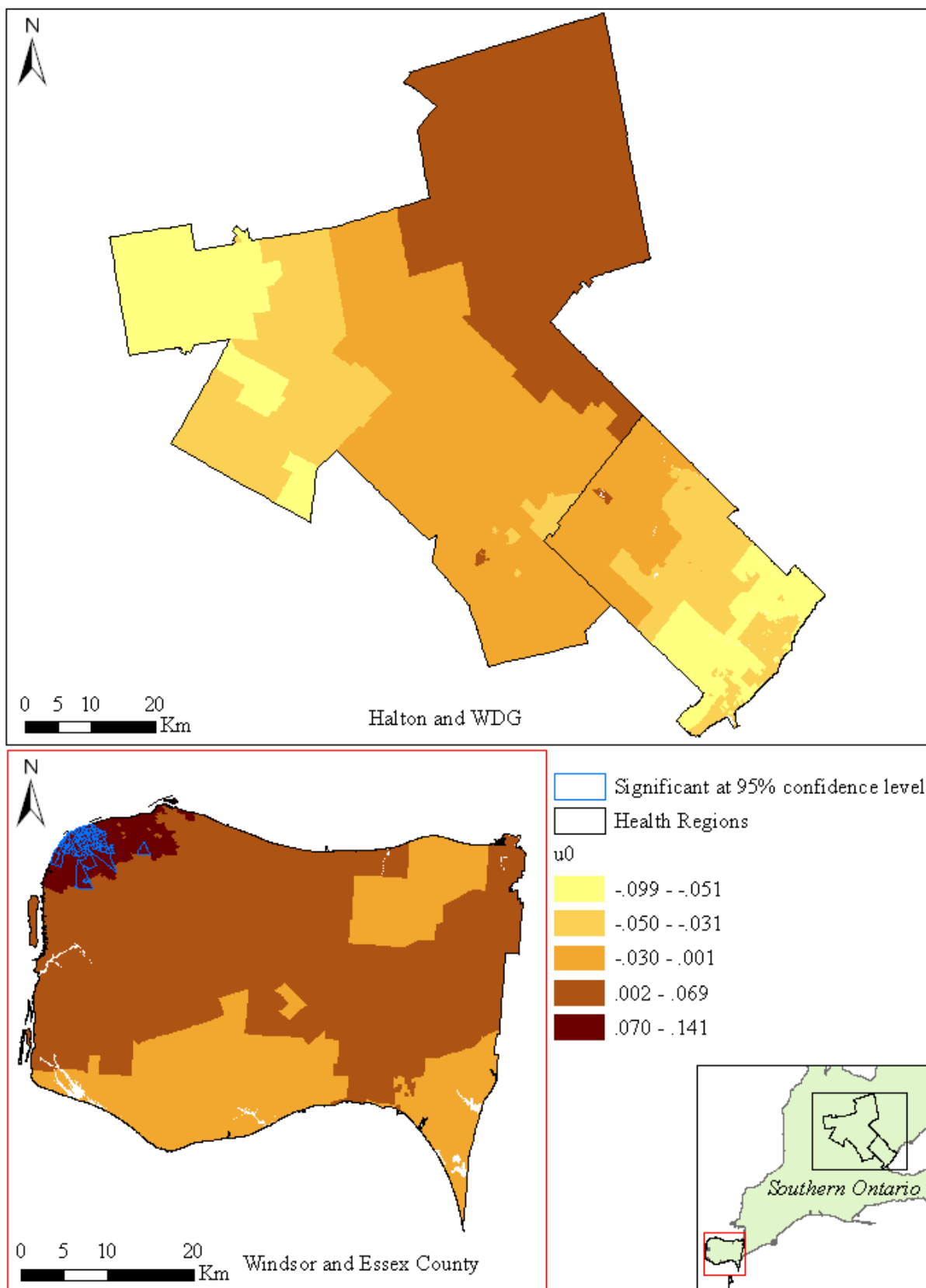


Figure G.11: Spatially correlated random effect  $u_0$  for LBW birth

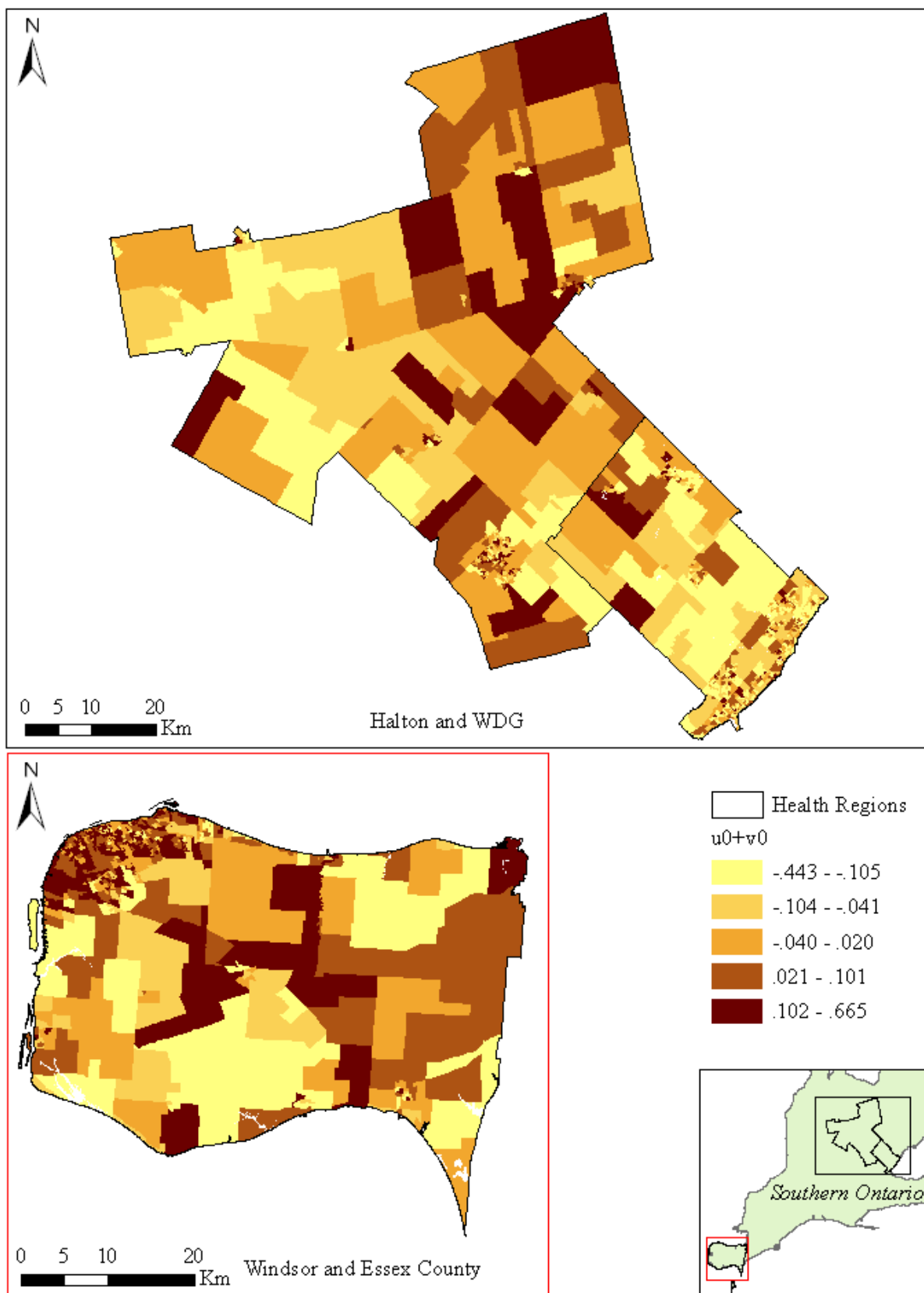


Figure G.12: Total random effects  $u_0 + v_0$  for LBW birth

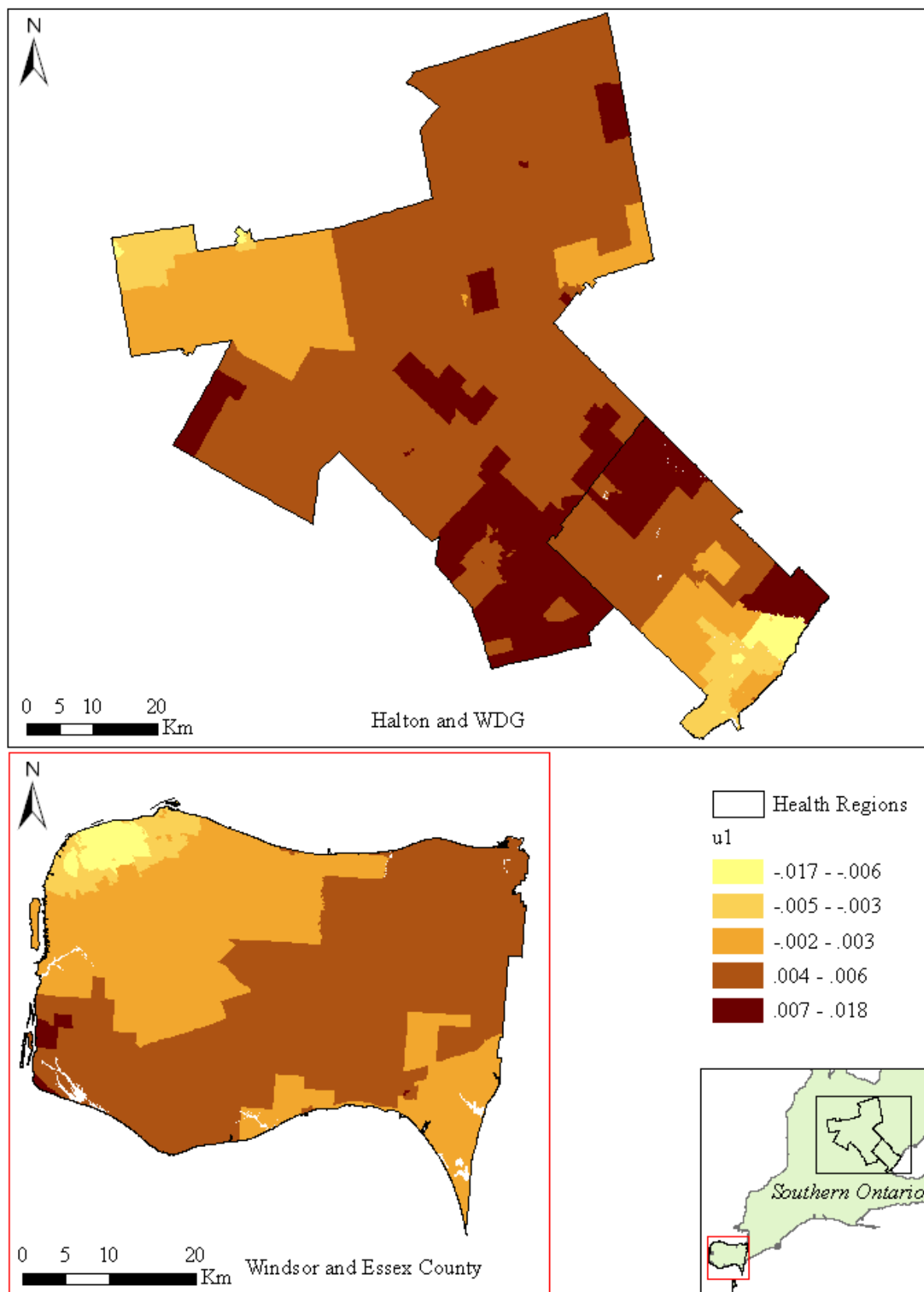


Figure G.13: Spatially correlated random effect of Factor 1 (behavioural)  $u_1$  for LBW birth

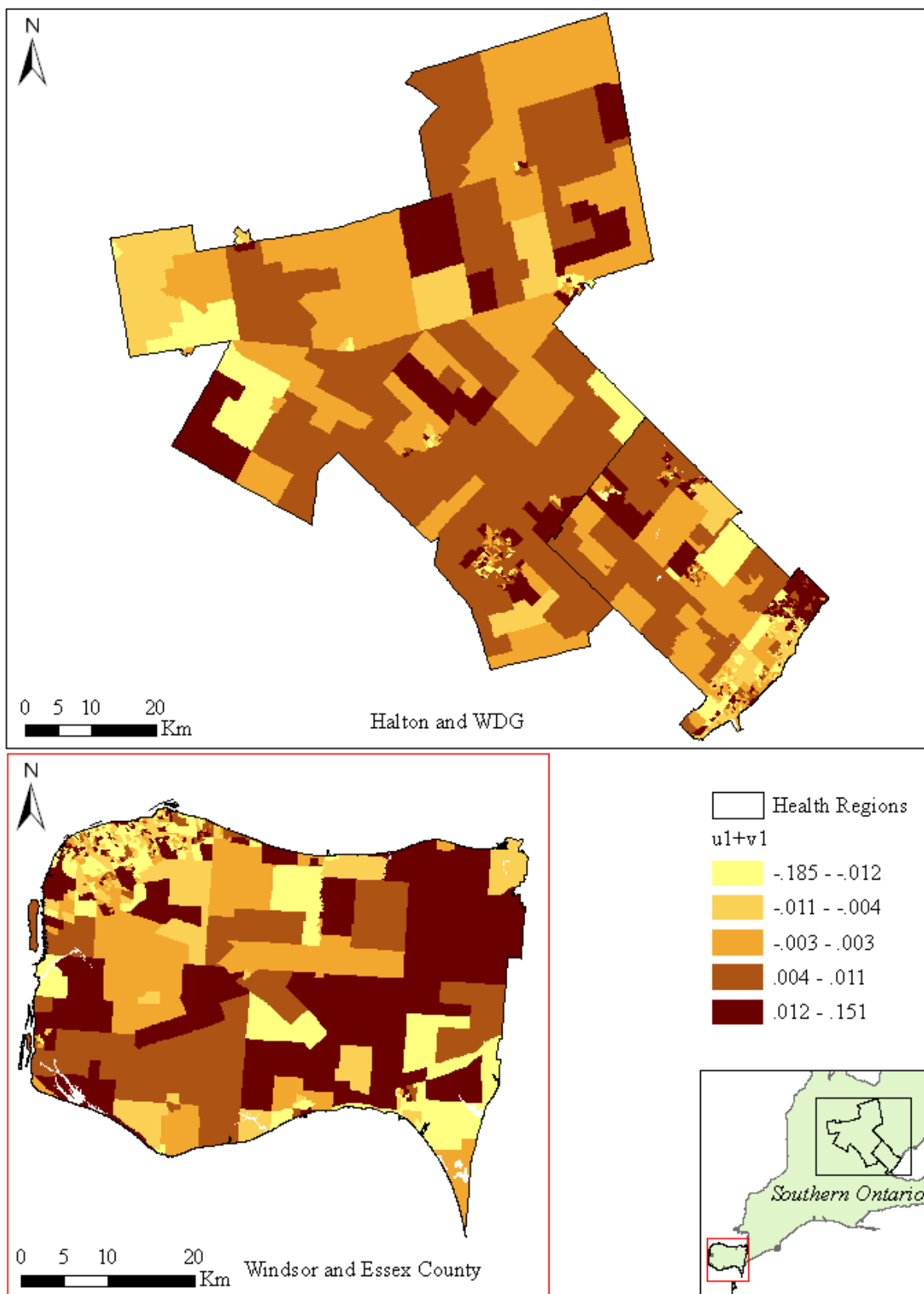


Figure G.14: Total random effects of Factor 1 (behavioural)  $u_l + v_l$  for LBW birth

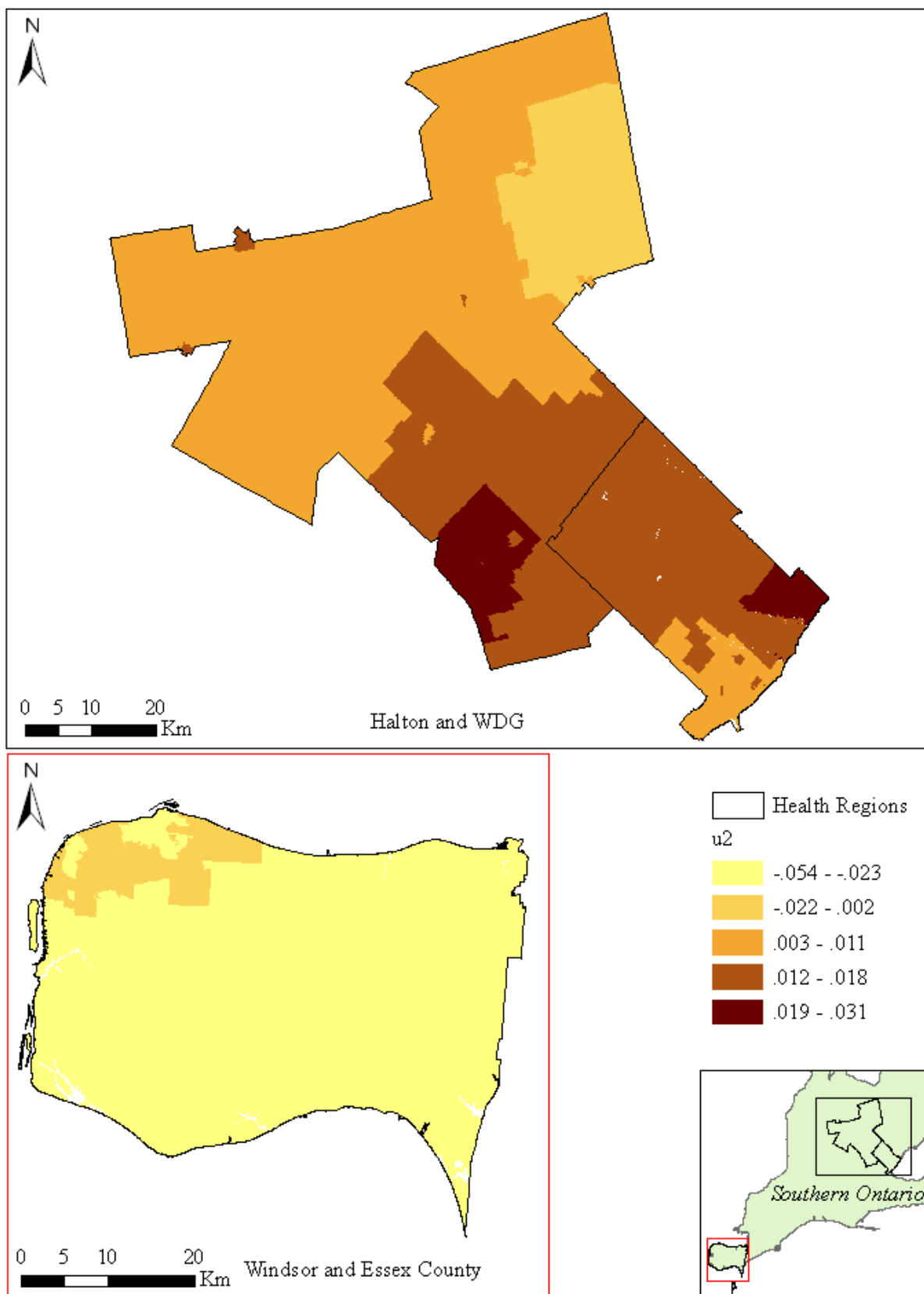


Figure G.15: Spatially correlated random effect of Factor 2 (SES)  $u_2$  for LBW birth

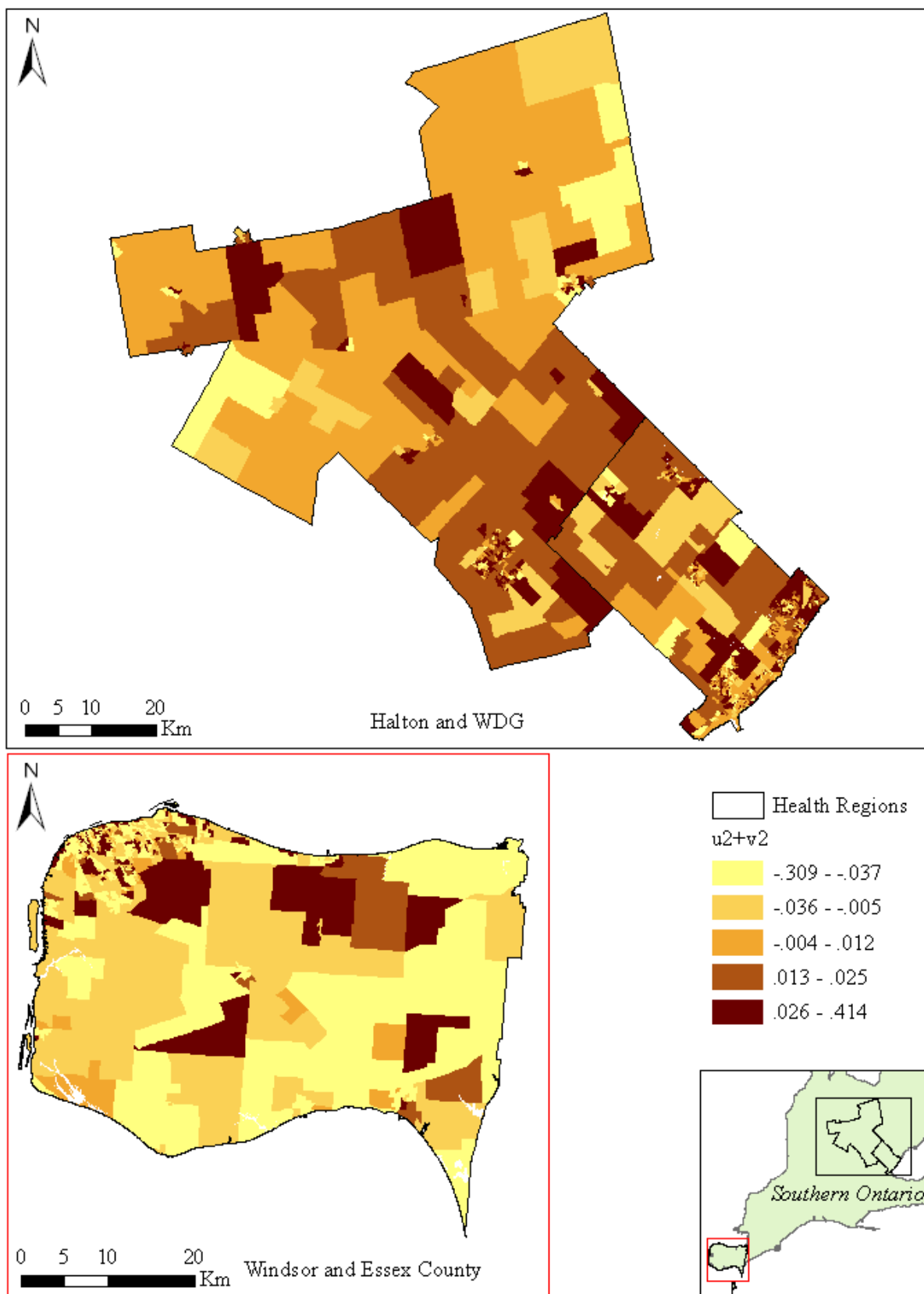


Figure G.16: Total random effects of Factor 2 (SES)  $u_2+v_2$  for LBW birth



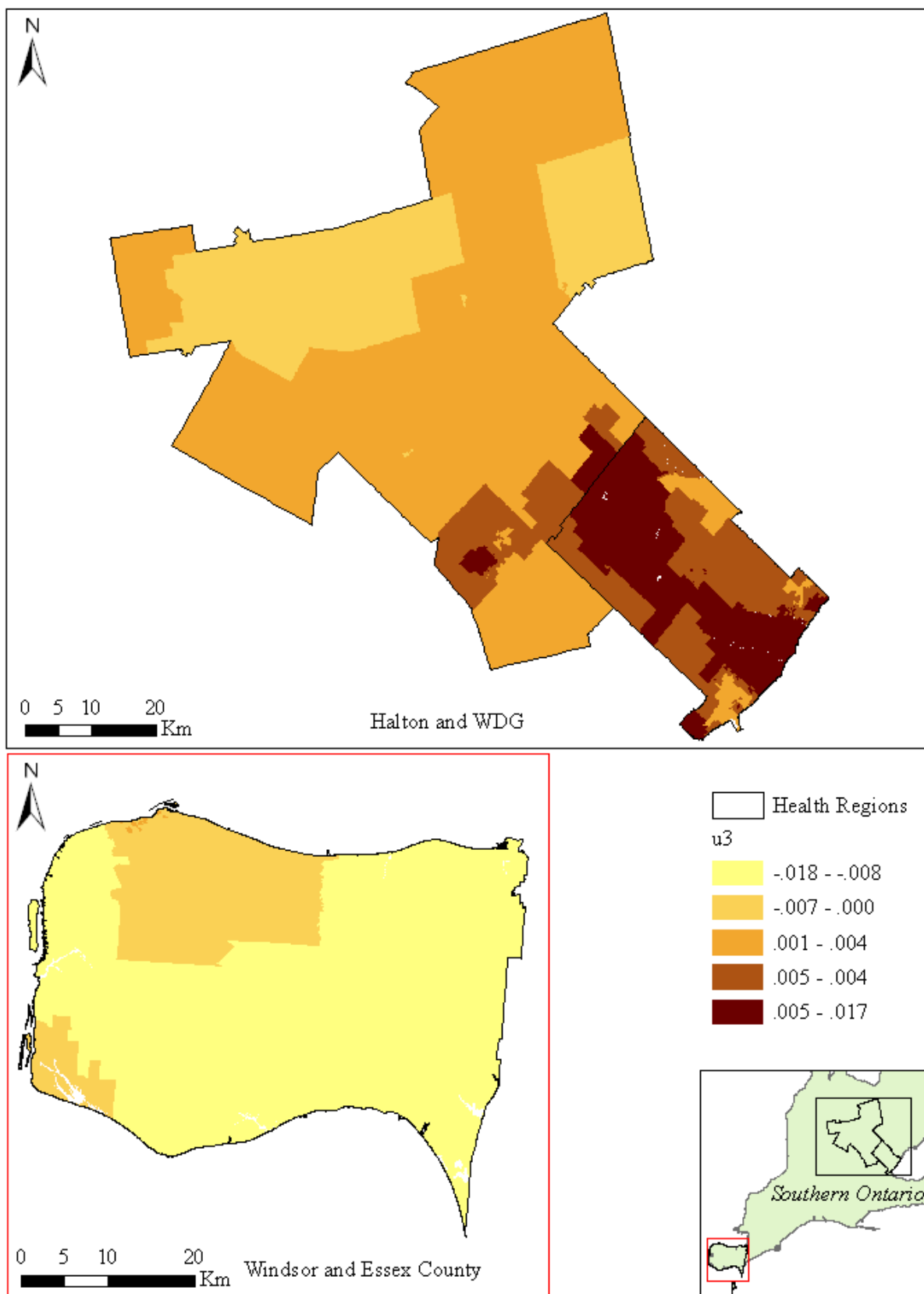


Figure G.17: Spatially correlated random effect of Factor 3 (psycho-social)  $u_3$  for LBW birth

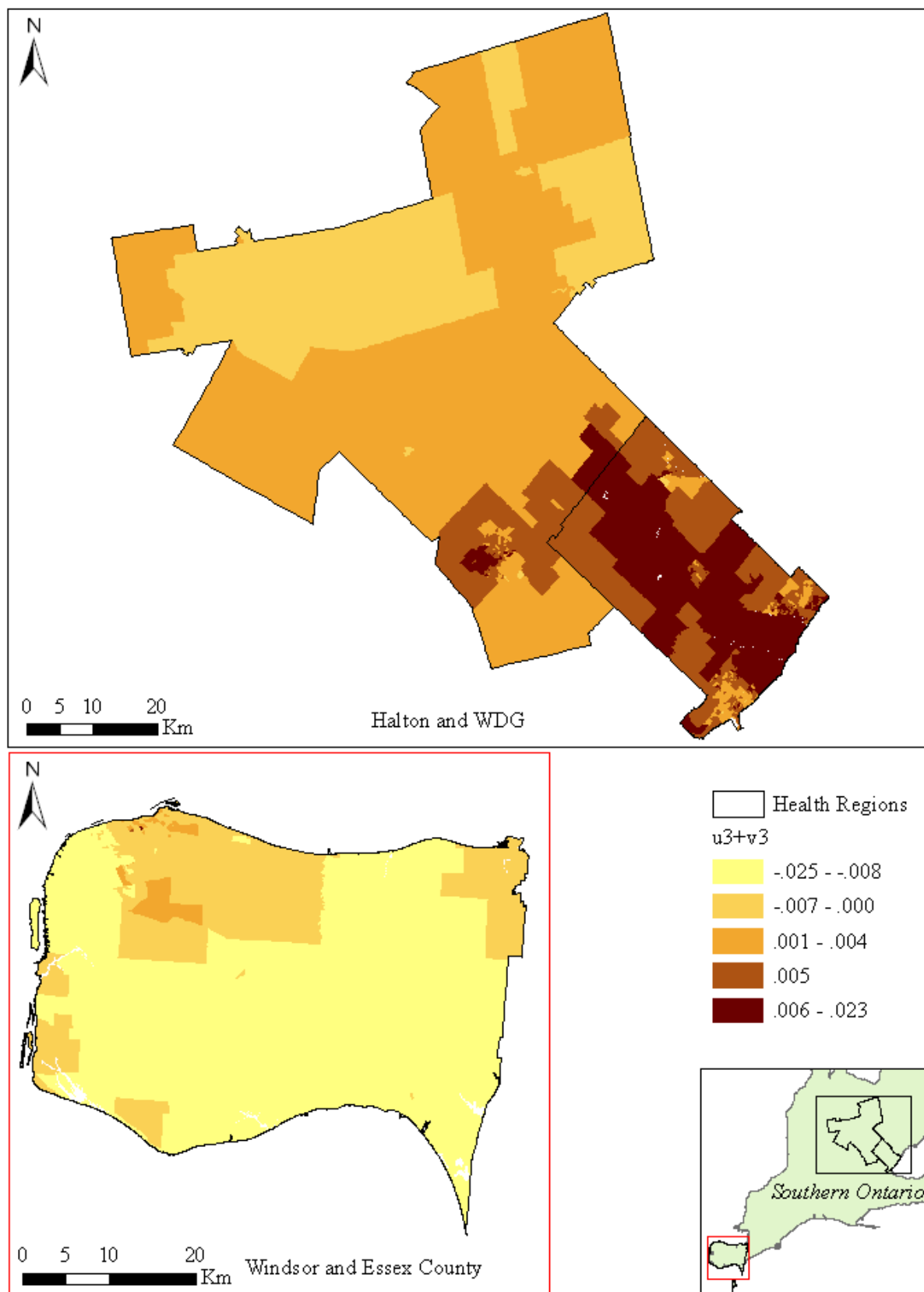


Figure G.18: Total random effects of Factor 3 (psycho-social)  $u_3+v_3$  for LBW birth

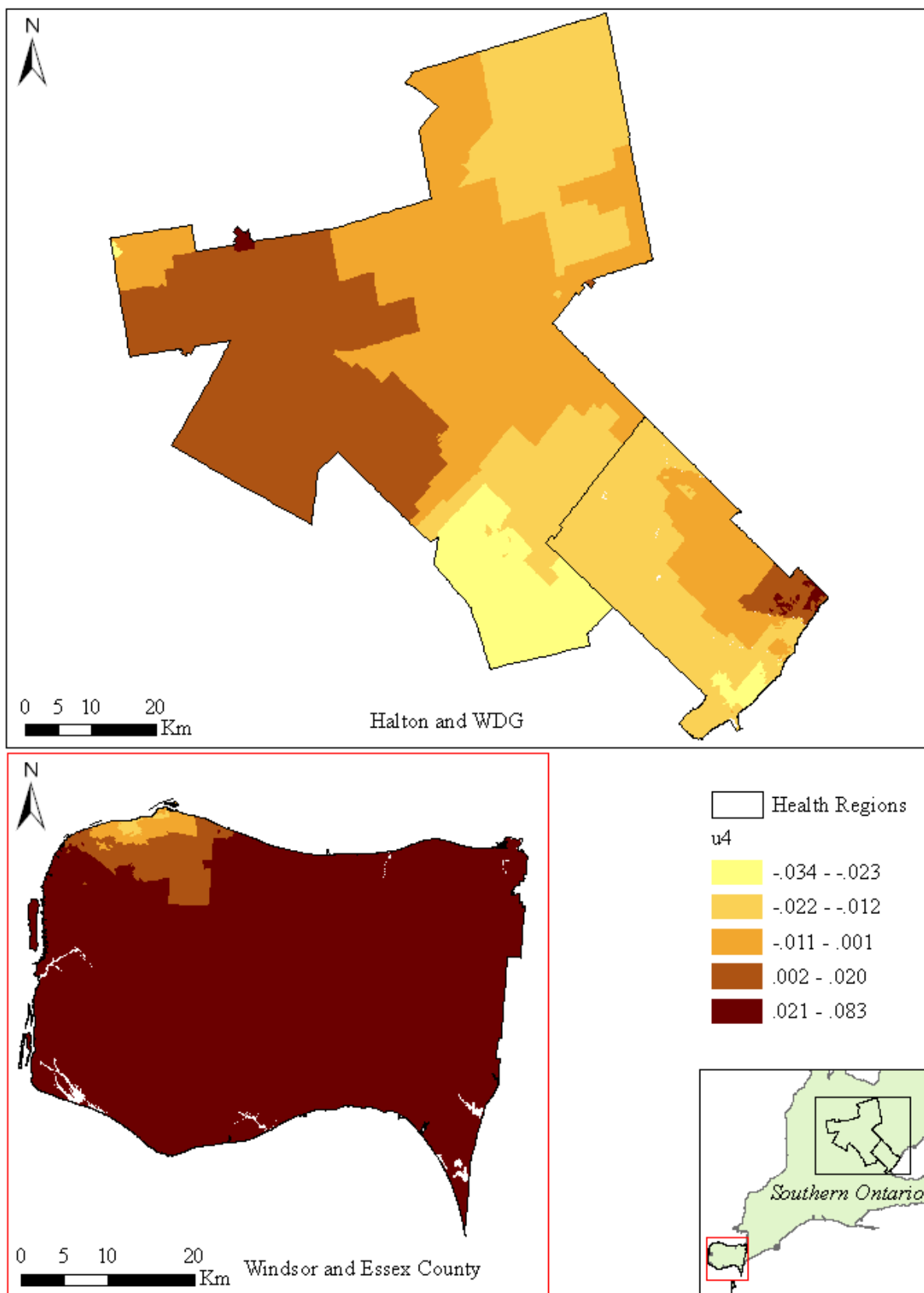


Figure G.19: Spatially correlated random effect of Factor 4 (health)  $u_4$  for LBW birth

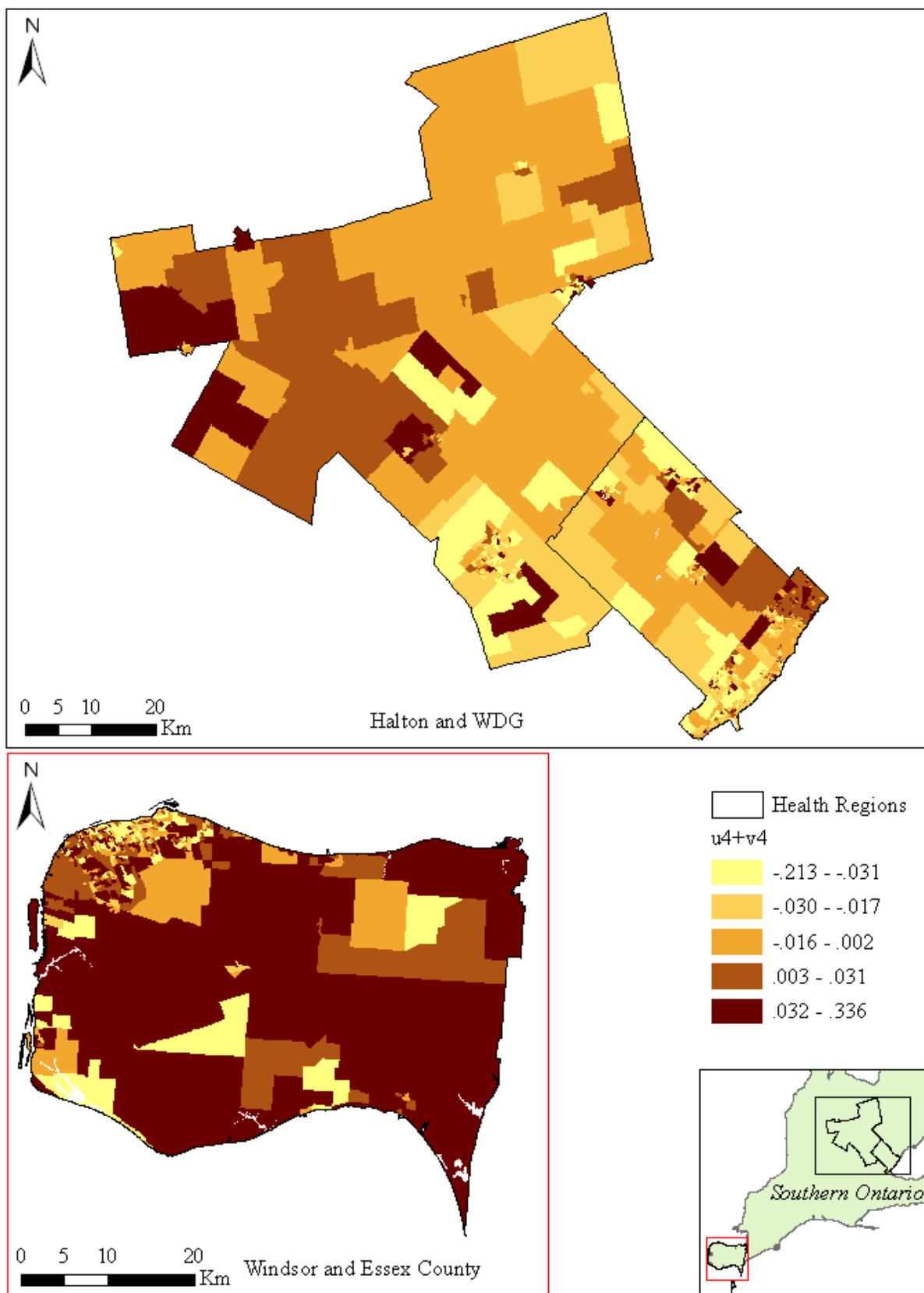


Figure G.20: Total random effects of Factor 4 (health)  $u_4 + v_4$  for LBW birth

## References

- Abel, E.L. and Hannigan, J.H. (1995). “Jshaped” relationship between drinking during pregnancy and birth weight: reanalysis of prospective epidemiological data. *Alcohol and Alcoholism*. Vol. 30, Issue 3, Pp. 345-355.
- Acito, F. and Anderson, R.D. (1980). A Monte Carlo comparison of factor analysis methods. *Journal of Marketing Research*. Vol. 17, Pp. 228-236.
- Addis, A., Moretti, M.E., Ahmed, S.F., Einarson, T.R., and Koren, G. (2001). Foetal effects of cocaine: an updated meta-analysis. *Reproductive Toxicology*. Vol. 15, Issue 4, Pp. 341-369.
- Alexander, G.R. (1998). Preterm birth: etiology, mechanisms and prevention. *Prenatal Neonatal Medicine*, Vol. 3, Pp. 3-9.
- Alexander, G.R. and Korenbrot, C.C. (1995). The role of prenatal care in preventing low birth weight. *Future Child*. Vol. 5, Issue 1, Pp. 103-120.
- Alkalay, A.L., Graham, J.M., and Pomerance, J.J. (1998). Evaluation of neonates born with intrauterine growth retardation: review and practice guidelines. *Journal of Perinatology*. Vol. 18, Issue 2, Pp. 142-151.
- American Academy of Pediatrics. (1997). Committee on Environmental Health. Noise: a hazard for the fetus and newborn. *Pediatrics*. Vol. 100, Issue 4, Pp. 724-727.
- American College of Obstetricians and Gynecologists. (1995). Preterm labor (No. 206) [ACOG technical bulletin]. *International Journal of Gynaecology and Obstetrics*, Vol. 50, Pp. 303-313.
- Amory, J.H., Adams, K.M., Lin, M.T., Hansen, J.A., Eschenbach, D.A., and Hitti, J. (2004). Adverse outcomes after preterm labor are associated with tumor necrosis factor-alpha polymorphism - 863, but not -308, in mother-infant pairs. *American Journal of Obstetrics and Gynecology*. Vol. 191, Issue 4, Pp. 1362-1367.
- Andersen, Hans Skifter (2003). *Urban Sores – On the interaction between segregation, urban decay and deprived neighbourhoods*. Ashgate, Aldershot, England.
- Andrews, W.W., Goldenberg, R.L., Hauth, J.C., Cliver, S.P., Copper, R., and Conner, M. (2006). Interconceptional antibiotics to prevent spontaneous preterm birth: a randomized clinical trial. *American Journal of Obstetrics and Gynecology*. Vol. 194, Issue 3, Pp. 617-623.
- Andres, R.L., and Day, M.C. (2000). Perinatal complications associated with maternal tobacco use. *Seminars in Neonatology*. Vol. 5, Issue 3, Pp. 231-241.
- Anotayanonth, S., Subhedar, N.V., Garner, P., Neilson, J.P., and Harigopal, S. (2004). Betamimetics for inhibiting preterm labour. *Cochrane Database of Systematic Reviews*. Issue 4. Art. No.:

CD004352. DOI: 10.1002/14651858.

- Anselin, L. (1995). Local indicators of spatial association: LISA. *Geographical Analysis*. Vol. 27, Issue 2, Pp. 93-116.
- Anselin, L. (2002). Under the hood: Issues in the specification and interpretation of spatial regression models. *Agricultural Economics*. Vol. 17, Issue 3, Pp. 247-67.
- Arbuckle, T. E., and Sherman, G. J. (1989). An analysis of birth weight by gestational age in Canada. *Canadian Medical Association Journal*, Vol. 140, Pp. 157-165.
- Armitage, P., and Colton, T. (1998). *Encyclopedia of biostatistics*. New York: J. Wiley.
- Armson, B.A., Dodds, L., Cervin, C., Christie-Haliburton, S., and Rinaldo, K. (2001). A preterm birth prevention project in Nova Scotia, Canada. *Maternal and Child Health Journal*. Vol. 5, Issue 3, Pp. 189-197.
- Armson, B.A., Dodds, L., Haliburton, S.C., Cervin, C., and Rinaldo, K. (2003). Impact of participation in the Halifax County Preterm Birth Prevention Project. *Journal of Obstetrics and Gynaecology*. Vol.25, Issue 3, Pp. 209-217.
- Armstrong, P., Armstrong, H., and David, C. (Eds.) (2001). *Unhealthy Times: political economy perspectives on health and care in Canada*. Oxford: Oxford University Press.
- Auger, N., Daniel, M., Platt, R.W., Wu, Y., Luo, Z.C., and Choinière, R. (2008). Association between perceived security of the neighbourhood and small-for-gestational-age birth. *Paediatric and Perinatal Epidemiology*. Vol. 22, Issue 5, Pp. 467-477.
- Bachmaier, M and Backes, M. (2008). Variogram or Semivariogram - Explaining the Variances in a Variogram. Article DOI: 10.1007/s11119-008-9056-2, *Precision Agriculture*. Springer-Verlag, Berlin, Heidelberg, New York.
- Balzi, D., Geddes, M., Brancker, A. and Parkin, D. (1995). Cancer Mortality in Italian Migrants and their offspring in Canada. *Cancer Causes and Control*. Vol. 6, Issue 1, Pp. 68-74.
- Barker, D.J., Gluckman, P.D., Godfrey, K.M., Harding, J.E., Owens, J.A., and Robinson, J.S. (1993). Foetal nutrition and cardiovascular disease in adult life. *Lancet*. Vol. 341 (8850), Pp. 938-941.
- Barker, D.J., Winter, P.D., Osmond, C., Margetts, B., and Simmonds, S.J. (1989). Weight in infancy and death from ischaemic heart disease. *Lancet*. Vol. 2 (8663), Pp. 577-580.
- Baron, R. M., and Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic and statistical considerations. *Journal of Personality and Social Psychology*. Vol. 51, Pp. 1173-1182.
- Barrett, F. (2000). Disease and Geography: the History of an Idea. *Geographical Monographs*. Vol. 23. Toronto, Atkinson College, York University.
- Bartha, J.L., Martinez-Del-Fresno, P., and Comino-Delgado, R. (2003). Early diagnosis of gestational

- diabetes mellitus and prevention of diabetes-related complications. *European Journal of Obstetrics, Gynecology, and Reproductive Biology*. Vol. 109, Issue 1, Pp.41-44.
- Bartholomew, D.J. (1987). *Latent Variable Models and Factor Analysis*. New York: Oxford University Press.
- Bartley, M. (2004). *Health Inequality: An Introduction to Theories, Concepts and Methods*. Cambridge: Polity Press.
- Beck, S., Wojdyla, D., Say, L., et al. (2010). The worldwide incidence of preterm birth: a systematic review of maternal mortality and morbidity. *Bulletin of the World Health Organization*. Vol. 88, Number 1, Pp. 31-38.
- Ben-Shlomo, Y. and Davey Smith, D. (2002). A life-course approach to chronic disease epidemiology: conceptual models, empirical challenges, and interdisciplinary perspectives. *International Journal of Epidemiology*. Vol. 31, Pp. 285-293.
- Bernstein, P.S., and Divon, M.Y. (1997). Etiologies of foetal growth restriction. *Clinical Obstetrics and Gynecology*. Vol. 40, Issue 4, Pp. 723-729.
- Berridge, V. and Blume, S. (2003). *Poor Health: Social inequality before and after the Black Report*. London: Frank Cass Publishers.
- Besag, J. (1989). Towards Bayesian image analysis. *Journal of Applied Statistics*. Vol. 16, Pp. 395-407.
- Besag, J., York, J. and Mollie, A. (1991). Bayesian image restoration, with two applications in spatial statistics. *Annals of the Institute of Statistical Mathematics*. Vol. 43, Pp. 1-59.
- Besag, J., and Newell, J. (1991). The detection of clusters in rare diseases. *Journal of the Royal Statistical Society, Series A*. Vol. 154, Pp. 143-155.
- Besag, J. and Kooperberg, C.L. (1995). On conditional and intrinsic autoregressions. *Biometrika*, Vol. 82, Pp. 733-746.
- Best Start. (2002). *Making a difference – program planning and implementation guide*. A Collaborative Project of Best Start: Maternal, Newborn & Early Child Development Resource Centre, The Perinatal Partnership Program of Eastern and Southeastern Ontario, and The Society of Obstetricians and Gynaecologists of Canada.
- Black, R.E. (2001). Micronutrients in pregnancy. *The British Journal of Nutrition*. Vol. 85 (Suppl 2), Pp. S193-S197.
- Blane, D., Bartley, M., and Davey Smith, G. (1998). Disease aetiology and materialist explanations of socioeconomic mortality differentials. *European Journal of Public Health*. Vol. 7, Pp. 385-391.
- Blane, D. (2006). The life course, the social gradient, and health. In Marmot, M., and Wilkinson, R.G (eds.). *Social Determinants of Health*. Second edition. New York: Oxford University Press.

- Blaxter, M. (1990). *Health and Lifestyles*. London: Tavistock.
- Blondel, B., and Bréart, G. (1995). Home visits during pregnancy: consequences on pregnancy outcome, use of health services, and women's situations. *Seminars in Perinatology*. Vol. 19, Issue 4, Pp. 263-271.
- Blumenfeld, Z., Lowenstein, L., Brook, O., Gonen, R., Ophir, E., and Samueloff, A. (2006). Does maternal height affect triplets' birth weight? *Medical Science Monitor: International Medical Journal of Experimental and Clinical Research*. Vol. 12, Issue 1, Pp. CR40-43.
- Bobak, M. (2000). Outdoor air pollution, low birth weight, and prematurity. *Environmental Health Perspectives*. Vol. 108, Pp. 173-176.
- Bodnar, L.M., Catov, J.M., Klebanoff, M.A., Ness, R.B., and Roberts, J.M. (2007). Prepregnancy body mass index and the occurrence of severe hypertensive disorders of pregnancy. *Epidemiology*. Vol. 18, Issue 2, Pp. 234-239.
- Bosma, H., Peter, R., Siegrist, J., and Marmot, M. (1998). Two alternative job stress models and the risk of coronary heart disease. *American Journal of Public Health*. Vol. 88, Issue 1, Pp. 68-74.
- Bourdieu, P. (1984). *Distinction*. London: Routledge.
- Bove, F., Shim, Y., and Zeitz, P. (2002). Drinking water contaminants and adverse pregnancy outcomes: a review. *Environmental Health Perspectives*. Vol. 110 (Suppl 1), Pp. 61-74.
- Breslow, N. E. and Clayton, D. G. (1993). Approximate inference in generalized linear mixed models. *Journal of the American Statistical Association*. Vol. 88, Pp. 9-25.
- Brocklehurst, P., and French, R. (1998). The association between maternal HIV infection and perinatal outcome: a systematic review of the literature and meta-analysis. *British Journal of Obstetrics and Gynaecology*. Vol. 105, Issue 8, Pp.836-848.
- Brooks, S.P. and Gelman, A. (1998) Alternative methods for monitoring convergence of iterative simulations. *Journal of Computational and Graphical Statistics*. Vol. 7, Pp. 434-455.
- Browne, M.L. (2006). Maternal exposure to caffeine and risk of congenital anomalies: a systematic review. *Epidemiology*. Vol. 17, Issue 3, Pp. 324-331.
- Browne, M.W. (1968). A comparison of factor analysis techniques. *Psychometrika*. Vol. 33, Pp. 267-334.
- Brunner, E.J. (1997). Stress and the biology of inequality. *British Medical Journal*. Vol. 314, Pp. 1472-1476.
- Brunner, E. and Marmot, M.G. (2006). Social organization, stress, and health. In Marmot, M.G. and Wilkinson, R.G. (eds.). *Social Determinants of Health*. 2<sup>nd</sup> edition. Oxford: Oxford University Press.
- Bucher, H.C., and Ragland, D.R. (1995). Socioeconomic indicators and mortality from coronary heart



- disease and cancer: a 22-year follow-up of middle-aged men. *American Journal of Public Health*. Vol. 85, Pp. 1231-1236.
- Bui, Q.Q., Tran, M.B., and West, W.L. (1986). A comparative study of the reproductive effects of methadone and benzo[a]pyrene in the pregnant and pseudopregnant rat. *Toxicology*. Vol. 42, Issue 2-3, Pp. 195-204.
- Buka, S.L., Brennan, R.T., Rich-Edwards, J.W., Raudenbush, S.W., and Earls, F. (2003). Neighborhood support and the birth weight of urban infants. *American Journal of Epidemiology*. Vol. 157, Issue 1, Pp. 1-8.
- Bukowski, J.A. (2004). Association between Air Pollution and Adverse Pregnancy Outcomes in Vancouver. *Environmental Health Perspectives*. Vol. 112, Issue 14, Pp. A792-794.
- Busacca, M., Balconi, G., Pietra, A., Vergara-Dauden, M., de Gaetano, G., and Dejana, E. (1984). Maternal smoking and prostacyclin production by cultured endothelial cells from umbilical arteries. *American Journal of Obstetrics and Gynecology*. Vol. 148, Issue 8, Pp. 1127-1130.
- Butler, R. and Bowlby, S. (1997). Bodies and spaces: an exploration of disabled people's experiences of public space. *Environment and Planning D: Society and Space*. Vol. 15, Pp. 411-433.
- Butte, N.F., Ellis, K.J., Wong, W.W., Hopkinson, J.M., and Smith, E.O. (2003). Composition of gestational weight gain impacts maternal fat retention and infant birth weight. *American Journal of Obstetrics and Gynecology*. Vol. 189, Issue 5, Pp. 1423-1432.
- Byrne, D. (1997). Social exclusion and capitalism: the reserve army across time and space. *Critical Social Policy*. Vol. 17, Pp. 27-51.
- Campbell, C. and Murray, M. (2004). Community health psychology: promoting analysis and action for social changes. *Journal of Health Psychology*. Vol. 9, Pp. 187-195.
- Campbell, J., Torres, S., Ryan, J., King, C., Campbell, D.W., Stallings, R.Y., et al. (1999). Physical and nonphysical partner abuse and other risk factors for low birth weight among full term and preterm babies: a multiethnic case-control study. *American Journal of Epidemiology*. Vol. 150, Issue 7, Pp. 714-726.
- Canada Institute for Health Information (2005). *Exploring the 70/30 split: how Canada's health care system is financed*. ISBN 1-55392-655-2. Ottawa, Ontario: CIHI.
- Canadian Institute for Health Information (2007). *Giving Birth in Canada: Regional Trends From 2001-2002 to 2005-2006. Analysis in Brief*. Ottawa, Ontario: CIHI.
- Canadian Institute for Health Information (2009). *Too Early, Too Small: A Profile of Small Babies Across Canada*. Ottawa, Ontario: CIHI.
- Carmichael, S.L., and Abrams, B. (1997). A critical review of the relationship between gestational weight gain and preterm delivery. *Obstetrics and Gynecology*. Vol. 89 (5 Pt 2), Pp. 865-873.

- Carter, H. (1981). *The Study of Urban Geography*, 3<sup>rd</sup> edition. London: Arnold.
- Casetti, E. (1997). The expansion method, mathematical modeling, and spatial econometrics. *International Regional Science Review*. Vol. 20, Pp. 9-32.
- Chan, E.C., Smith, R., Lewin, T., Brinsmead, M.W., Zhang, H.P., Cubis, J., et al. (1993). Plasma corticotropin-releasing hormone, beta-endorphin and cortisol inter-relationships during human pregnancy. *Acta Endocrinologica*. Vol. 128, Issue 4, Pp. 339-344.
- Cherniak, D., Grant, L., Mason, R., Moore, B., and Pellizzari, R. (2005). Society of Obstetricians and Gynaecologists of Canada. Intimate partner violence consensus statement. *Journal of Obstetrics and Gynaecology Canada*. Vol. 27, Issue 4, Pp. 365- 418.
- Chia, S.E., Lee, J., Chia, K.S., and Chan, O.Y. (2004). Low birth weight in relation to parental occupations-a population-based registry in Singapore (1994-1998). *Neurotoxicology and Teratology*. Vol. 26, Issue 2, Pp. 285-290.
- Clarren, S. (1988). Alcohol related birth defects: The clinical situation as defined over 15 years of experience. *Proceedings Alcohol and Child/Family Health*. Vancouver.
- Clayton, D.G. and Kaldor, J. (1987). Empirical Bayes estimates of age-standardized relative risks for use in disease mapping. *Biometrics*. Vol. 43, Pp. 671-681.
- Cliff, A. and Haggett, P. (1988). *Atlas of Disease Distributions: analytic approaches to epidemiological data*. Oxford, Blackwell.
- Coburn, D. (2000). Income inequality, lowered social cohesion, and the poorer health status of populations: the role of neo-liberalism. *Social Science and Medicine*. Vol. 51, Pp. 135-146.
- Congdon, P. (2000). A Bayesian approach to prediction using the gravity model, with an application to patient flow modelling. *Geographical Analysis*. Vol. 32, Issue 2, Pp. 205-224.
- Coomarasamy, A., Thangaratnam, S., Gee, H., Khan, K.S. (2006). Progesterone for the prevention of preterm birth: a critical evaluation of evidence. *European Journal of Obstetrics, Gynecology, and Reproductive Biology*. Vol. 129, Issue 2, Pp. 111-118.
- Congress for the New Urbanism. (1999). *Chapter of the New Urbanism*. McGraw-Hill, Inc.
- Cotch, M.F., Pastorek, J.G., Nugent, R.P., Hillier, S.L., Gibbs, R.S., Martin, D.H., et al. (1997). Trichomonas vaginalis associated with low birth weight and preterm delivery. The Vaginal Infections and Prematurity Study Group. *Sexually Transmitted Diseases*. Vol. 24, Issue 6, Pp. 353-360.
- Covington, D.L., Justason, B.J., and Wright, L.N. (2001). Severity, manifestations, and consequences of violence among pregnant adolescents. *The Journal of Adolescent Health*. Vol. 28, Issue 1, Pp. 55-61.
- Cowles, M.K. and Carlin, B. (1996). Markov chain Monte Carlo convergence diagnostics: a

- comparative review. *Journal of the American Statistical Association*. Vol. 91, Pp. 883-904.
- Creasy, R.K., Gummer, B.A., and Liggins, G.C. (1980). System for predicting spontaneous preterm birth. *Obstetrics and Gynecology*. Vol. 55, Issue 6, Pp. 692-695.
- Cressie, N.A. and Chan, N.H. (1989). Spatial modeling of regional variables. *Journal of the American Statistical Association*, Vol. 84, Pp. 393-401.
- Cressie, N. (1993). *Statistics for Spatial Data, revised edition*. Wiley, New York.
- Cromley, E. and McLafferty, S. (2002). *GIS and Public Health*. New York: Guilford.
- Crowther, C.A. (2001). Hospitalisation and bed rest for multiple pregnancy. *Cochrane Database of Systematic Reviews*. Issue 1. Art. No.: CD000110. DOI: 10.1002/14651858.
- Curtis, S. (2004). *Health and Inequality: Geographical Perspectives*. London: SAGE Publications
- Dandeker, C. (1990). *Surveillance, Power and Modernity*. Cambridge: Polity.
- Dannefer, D. (2003). Cumulative advantage/disadvantage and the life course: cross-fertilizing age and social science theory. *Journal of Gerontology*. Vol. 58B, Pp. 327-337.
- David, S. Andrew, T. Nicky, B. and Dave, L. (2003). *WinBUGS User Manual*. Version 1.4.
- Dear, M., Gaber, L., Takahashi, L. and Wilton, R. (1997). Seeing people differently: the socio-spatial construction of disability. *Environment and Planning D: Society and Space*. Vol. 15, Issue 4, Pp. 455-480.
- Dejmek, J., Selevan, S.G., Benes, I., Solansky, I., and Srám, R.J. (1999). Foetal growth and maternal exposure to particulate matter during pregnancy. *Environmental Health Perspectives*. Vol. 107, Pp. 475-480.
- Dizon-Townson, D.S. (2001). Preterm labour and delivery: a genetic predisposition. *Paediatric and Perinatal Epidemiology*. Vol. 15, Suppl. 2, Pp. 57-62.
- Dollfus, C. Paletta, M., Siegel, E. and Cross A.W. (1990). Infant mortality: a practical approach to the analysis of the leading causes of death and risk factors. *Pediatrics*. Vol. 86, Pp. 176-183.
- Doroshenko, A., Sherrard, J., and Pollard, A.J. (2006). Syphilis in pregnancy and the neonatal period. *International Journal of STD & AIDS*. Vol. 17, Issue 4, Pp. 221-227.
- Dreever, F., Whitehead, M. and Roden, M. (1996). Current patterns and trends in male mortality by social class (based on occupation). *Population Trends*. Vol. 86 (Winter), Pp. 15-20.
- Dunn, J.R., Veenstra, G., and Ross, N. (2006). Psychosocial and neo-material dimensions of SES and health revisited: Predictors of self-rated health in a Canadian national survey. *Social Science and Medicine*. Vol. 62, Pp. 1465-1473.
- Elfenbein, D.S., and Felice, M.E. (2003). Adolescent pregnancy. *Pediatric Clinics of North America*. Vol. 50, Issue 4, Pl. 781-800, viii.
- Elo, I.T., Culhane, J.F., Kohler, I.V., O'Campo, P., Burke, J.G., Messer, L.C., Kaufman, J.S., Laraia,

- B.A., Eyster, J., Holzman, C. (2009). Neighbourhood deprivation and small-for-gestational-age term births in the United States. *Paediatric and Perinatal Epidemiology*. Vol. 23, Issue 1, Pp. 87-96.
- Elstad, J.I. (1998). The psycho-social perspective on social inequalities in health. In Bartley, M., Blane, D., and Smith, G.D. (eds.). *The Sociology of Health Inequalities*. Oxford: Blackwell.
- Elter, K., Ay, E., Uyar, E., and Kavak, Z.N. (2004). Exposure to low outdoor temperature in the midtrimester is associated with low birth weight. *The Australian & New Zealand Journal of Obstetrics & Gynaecology*. Vol. 44, Issue 6, Pp. 553-557.
- Emslie, C., Hunt, K. and Macintyre, S. (1999). Problematising gender, work and health: the relationship between gender, occupational grade, working conditions and minor morbidity in full-time bank employees. *Social Science and Medicine*. Vol. 48, Issue 1, Pp. 33-48.
- Eslick, G.D., Yan, P., Xia, H.H., Murray, H., Spurrett, B., and Talley, N.J. (2002). Foetal intrauterine growth restrictions with *Helicobacter pylori* infection. *Alimentary Pharmacology & Therapeutics*. Vol. 16, Issue 9, Pp. 1677-1682.
- Farley, T.A., Mason, K., Rice, J., Habel, J.D., Scribner, R., and Cohen, D.A. (2006). The relationship between the neighbourhood environment and adverse birth outcomes. *Paediatric and Perinatal Epidemiology*. Vol. 20, Issue 3, Pp. 188-200.
- Fleiss, J. L. (1981). *Statistical Methods for Rates and Proportions*, 2nd edition. New York: Wiley.
- Fleiss, J. L. and Cuzick, J. (1979). The reliability of dichotomous judgements: Unequal numbers of judges per subject. *Applied Psychological Measurement*. Vol. 3, Pp. 537-542.
- Fotheringham, A.S., Brunson, C., and Charlton, M. (2002). *Geographically Weighted Regression: the analysis of spatially varying relationships*. Chichester: Wiley.
- Foucault, M. (1979). *Discipline and Punish*. Harmondsworth: Penguin.
- Fried, P.A., and Smith, A.M. (2001). A literature review of the consequences of prenatal marihuana exposure. An emerging theme of a deficiency in aspects of executive function. *Neurotoxicology and Teratology*. Vol.23, Issue 1, Pp. 1-11.
- Gatrell, A. (2002). *Geographies of Health*. Oxford: Blackwell.
- Gazmararian, J. A., Lazorick, S., Spitz, A. M., Ballard, T. J., Saltzman, L. E., and Marks, J. S. (1996). Prevalence of violence against pregnant women. *Journal of the American Medical Association*. Vol. 275, Pp. 1915-1920.
- Geary, R. C. (1954). The contiguity ratio and statistical mapping. *The Incorporated Statistician*. Vol. 5, Pp. 115-145.
- Gelman, A. and Rubin, D.B. (1992). Inference from iterative simulation using multiple sequence. *Statistical Science*. Vol. 7, Pp. 457-472.

- Generoso, W.M., Rutledge, J.C., Cain, K.T., Hughes, L.A., and Braden, P.W. (1987). Exposure of female mice to ethylene oxide within hours after mating leads to foetal malformation and death. *Mutation Research*. Vol. 176, Pp. 269-274.
- Geman, S. and Geman, D. (1984). Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. Vol. 6, Issue 6, Pp. 721-741.
- Getis, A. and J.K. Ord (1992). The analysis of spatial association by use of distance statistics. *Geographical Analysis*. Vol. 24, Issue 3, Pp. 189-206.
- Geweke, J. (1992). Evaluating the accuracy of sampling-based approaches to calculating posterior moments. In *Bayesian Statistics 4*, (ed. J. M. Bernardo, J. O. Berger, A. P. Dawid, and A. F. M. Smith). Clarendon Press, Oxford, UK.
- Gibbs, R.S., Romero, R., Hillier, S.L., Eschenbach, D.A., and Sweet, R.L. (1992). A review of premature birth and subclinical infection. *American Journal of Obstetrics and Gynecology*. Vol. 166, Issue 5, Pp.1515-1528.
- Giddens, A. (1984). *The Constitution of Society: Outline of a Theory of Structuration*. Cambridge: Polity.
- Gilbertson, M. and Brophy, J., (2001). Community Health Profile of Windsor, Ontario, Canada: Anatomy of a Great Lakes Area of Concern. *Environmental Health Perspective*. Vol. 109, Suppl. 6, Pp. 827-843.
- Gilks, W.R., Richardson, S., and Spiegelhalter, D.J. (1996). *Markov Chain Monte Carlo in Practice*. Chapman & Hall, London.
- Goldenberg, R.L., Hoffman, H.J., and Cliver, S.P. (1998). Neurodevelopmental outcome of small-for-gestational-age infants. *European Journal of Clinical Nutrition*. Vol. 52 supplement 1, S54-S58.
- Goulet, C., Gevry, H., Lemay, M., Gauthier, R.J., Lepage, L., Fraser, W., et al. (2001). A randomized clinical trial of care for women with preterm labour: home management versus hospital management. *Canadian Medical Association Journal*. Vol. 164, Issue 7, Pp. 985-991.
- Guttman, L. (1953). Image theory of the structure of quantitative variates. *Psychometrika*. Vol. 18, Pp. 277-296.
- Hack, M., Klein, N.K., and Taylor, H.G. (1995). Long-term developmental outcomes of low birth weight infants. *The Future of Children*. Vol. 5, Pp. 176-196.
- Hack, M. (1998). Effects of intrauterine growth retardation on mental performance and behaviour, outcomes during adolescence and adulthood. *European Journal of Clinical Nutrition*. Vol. 52, Supplement 1, S65-S71.
- Haider, B.A., and Bhutta, Z.A. (2006) Multiple-micronutrient supplementation for women during

- pregnancy. *Cochrane Database of Systematic Reviews*. Issue 4. Art. No.: CD004905. DOI: 10.1002/14651858.
- Haining, R. (2003). *Spatial Data Analysis - theory and practice*. Cambridge University Press, Cambridge, UK .
- Hanke, W., and Jurewicz, J. (2004). The risk of adverse reproductive and developmental disorders due to occupational pesticide exposure: an overview of current epidemiological evidence. *International Journal of Occupational Medicine and Environmental Health*. Vol. 17, Issue 2, Pp. 223-243.
- Hanson, D., Hanson, J., Vardon, P., McFarlane, K., Lloyd, J., Muller, R., et al. (2005). The injury iceberg: An ecological approach to planning sustainable community safety interventions. *Health Promotion Journal of Australia*. Vol. 16, Issue 1, Pp 5-10.
- Harding, D.J. (2003). Counterfactual models of neighborhood effects: the effect of neighborhood poverty on dropping out and teenage pregnancy. *American Journal of Sociology*. Vol. 109, Issue 3, Pp. 676-719.
- Harman, H.H. (1976). *Modern Factor Analysis*, 3rd ed. Chicago: University of Chicago Press.
- Harris, C.W. and Kaiser, H.F. (1964). Oblique factor analytic solutions by orthogonal transformations. *Psychometrika*. Vol. 29, Pp. 347-362.
- Hatch, M.C., Shu, X.O., McLean, D.E., Levin, B., Begg, M., Reuss, L., et al. (1993). Maternal exercise during pregnancy, physical fitness, and foetal growth. *American Journal of Epidemiology*. Vol. 137, Issue 10, Pp. 1105-1114.
- Hatch, M., Levin, B., Shu, X.O., and Susser, M. (1998). Maternal leisure-time exercise and timely delivery. *American Journal of Public Health*. Vol. 88, Issue 10, Pp. 1528-1533.
- Heaman, M.I. (2005). Relationship between physical abuse during pregnancy and risk factors for preterm birth among women in Manitoba. *Journal of Obstetric, Gynecologic, and Neonatal Nursing*. Vol. 34, Pp. 721-731.
- Heath, D. (1983). The maturing person. In Walsh, R.N. and Shapiro, D.H. *Beyond Health and Normality. Explorations of exceptional psychological health*. New York: Van Nostrand.
- Hebel, J., Fox, N. and Sexton, M. (1988). Dose-response of birth weight to various measures of maternal smoking during pregnancy. *Journal of Clinical Epidemiology*. Vol. 41, Issue 5, Pp. 483-489.
- Heidelberger, P. and Welch, P.D. (1983). Simulation run length control in the presence of an initial transient . *Operations Research*. Vol. 31, Pp. 1109-1144.
- Helman C.G. (2000). *Culture, Health and Illness*. 4th Ed. Oxford: Butterworth-Heinemanne.
- Hendrickson, A.E. and White, P.O. (1964). PROMAX: a quick method for rotation to oblique simple

- structure. *British Journal of Statistical Psychology*. Vol. 17, Pp. 65-70.
- Hendrix, N., and Berghella, V. (2008). Non-placental causes of intrauterine growth restriction. *Seminars in Perinatology* Vol. 32, Issue 3, Pp. 161-165.
- Henley, S. (1981). *Nonparametric Geostatistics*. London: Applied Science Publishers, Ltd.
- Hertzman, C., Power, C. Matthews, S. and Manor, O. (2001). Using an interactive framework of society and life course to explain self-rated health in early adulthood. *Social and Scientific Medicine*. Vol. 53, Pp. 1575-85.
- Hillier, S.L., Nugent, R.P., Eschenbach, D.A., Krohn, M.A., Gibbs, R.S., Martin, D.H., et al. (1995). Association between bacterial vaginosis and preterm delivery of a low birth weight infant. The Vaginal Infections and Prematurity Study Group. *The New England Journal of Medicine*. Vol. 333, Issue 26, Pp. 1737-1742.
- Hobel C.J., Dunke-Schetter C., and Roesch S. (1998). Maternal stress as a signal to foetus. *Prenatal and Neonatal Medicine*. Vol. 3, Pp. 116-120.
- Hodnett, E.D., and Fredericks, S. (2003). Support during pregnancy for women at increased risk of low birthweight babies. *Cochrane Database of Systematic Reviews*. Issue 3. Art. No.: CD000198. DOI: 10.1002/14651858.
- Hogue, C.J., Hoffman, S., and Hatch, M.C. (2001). Stress and preterm delivery: a conceptual framework. *Paediatric and Perinatal Epidemiology*. Vol.15 (Supplement), Pp. 30-40.
- Holzman, C., and Paneth, N. (1994). Maternal cocaine use during pregnancy and perinatal outcomes. *Epidemiological Reviews*. Vol.16, Issue 2, Pp. 315-334.
- Holzman, C., Bullen, B., Fisher, R., Paneth, N., and Reuss, L. (2001). Pregnancy outcomes and community health: the POUCH study of preterm delivery. *Paediatric and Perinatal Epidemiology*. Vol. 15 (Supplement 2), Pp. 136-158.
- Honest, H., Bachmann, L.M., Sundaram, R., Gupta, J.K., Kleijnen, J., and Khan, K.S. (2004). The accuracy of risk scores in predicting preterm birth - a systematic review. *Journal of Obstetrics and Gynaecology*. Vol. 24, Issue 4, Pp. 343-359.
- Honest, H., Bachmann, L.M., Ngai, C., Gupta, J.K., Kleijnen, J., and Khan, K.S. (2005). The accuracy of maternal anthropometry measurements as predictor for spontaneous preterm birth - a systematic review. *European Journal of Obstetrics, Gynecology, and Reproductive Biology*. Vol. 119, Issue 1, Pp. 11-20.
- Hotelling, H. (1933). Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology*. Vol. 24 Pp. 417-441, 498-520.
- Hox, J. (2002). *Multilevel analysis*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Hueston, W.J., Knox, M.A., Eilers, G., Pauwels, J., and Lonsdorf, D. (1995). The effectiveness of

- preterm-birth prevention educational programs for high-risk women: a metaanalysis. *Obstetrics and Gynecology*. Vol. 86, (4 Pt 2), Pp. 705-712.
- Hulse, G.K., Milne, E., English, D.R., and Holman, C.D. (1997). The relationship between maternal use of heroin and methadone and infant birth weight. *Addiction*. Vol.92, Issue 11, Pp. 1571-1579.
- Ickovics, J.R., Kershaw, T.S., Westdahl, C., Rising, S.S., Klima, C., Reynolds, H., et al. (2003). Group prenatal care and preterm birth weight: results from a matched cohort study at public clinics. *Obstetrics and Gynecology*. Vol. 102, (5 Pt 1), Pp. 1051-1057.
- Institute of Health Economics (2008). *IHE Report-Determinants and Prevention of Low Birth Weight: A Synopsis of the Evidence*. Alberta, Canada.
- Jarvis, M.J. and Wardle, J. (2006). Social patterning of individual health behaviours: the case of cigarette smoking. In Marmot, M., and Wilkinson, R.G (eds.). *Social Determinants of Health*. 2<sup>nd</sup> edition. New York: Oxford University Press.
- Jeannine, C. (Ed.) (2010). *Social and Behavioural Foundations of Public Health*. SAGE Publications, Inc.
- Jerret, M., Burnett, R., Kanaroglou, P., Eyles, J., Finkelstein, N., Giovis, C. and Brook, J. (2001). A GIS: environmental justice analysis of particulate air pollution in Hamilton, Canada. *Environment and Planning A*. Vol. 33, Issue 6, Pp. 955-973.
- Johnson, J.W., and Yancey, M.K. (1996). A critique of the new recommendations for weight gain in pregnancy. *American Journal of Obstetrics and Gynecology*. Vol. 174 (1 Pt 1), Pp.254-258.
- Joreskog, K.G. (1977). Factor analysis by least squares and maximum likelihood methods. In *Statistical Methods of Digital Computers, Vol. III of Mathematical Methods for Digital Computers*, edited by Enslein, K., Ralston, A., and Wilf, H.S. New York: John Wiley & Sons.
- Joseph, K.S. (1998). *Preterm birth in Canada*. Background paper prepared for the Preterm Birth Prevention Conference, Ottawa, Ontario, Canada.
- Joseph, M. and Chaskin, R. (2010) Living in a Mixed-Income Development: Resident Perceptions of the Benefits and Disadvantages of Two Developments in Chicago. *Urban Studies*. Vol. 47, Pp. 2347-2366.
- Judd, C. M., and Kenny, D. A. (1981). Process analysis: Estimating mediation in treatment evaluations. *Evaluation Review*. Vol. 5, Pp. 602-619.
- Kaiser, H.F. (1958). The varimax criterion for analytic rotation in factor analysis. *Psychometrika*. Vol. 23, Pp. 187-200.
- Kaiser, H.F. and Caffrey, J. (1965). Alpha factor analysis. *Psychometrika*. Vol. 30, Pp.1-14.
- Karlin, S., Cameron, P. E., and Williams, P. (1981). Sibling and parent-offspring correlation with variable family age. *Proceedings of the National Academy of Science, U.S.A.* Vol. 78, Pp. 2664-



- Kawachi, I., Kennedy, B.P., and Glass, R. (1999). Social capital and self-rated health: A contextual analysis. *American Journal of Public Health*. Vol. 89, Pp. 1187-1193.
- Kawachi, I. and Berkman, L.F. (Eds.). (2000). *Social Epidemiology*. New York: Oxford University Press.
- Kawachi, I. and Berkman, L.F. (Eds.). (2003). *Neighbourhoods and Health*. New York: Oxford University Press.
- Keen, C.L., Clegg, M.S., Hanna, L.A., Lanoue, L., Rogers, J.M., Daston, G.P., et al. (2003). The plausibility of micronutrient deficiencies being a significant contributing factor to the occurrence of pregnancy complications. *The Journal of Nutrition*. Vol. 133, Issue 5, Suppl 2, Pp. 1597S-1605S.
- Kenyon, S., Boulvain, M., and Neilson, J.P. (2004). Antibiotics for preterm rupture of the membranes: a systematic review. *Obstetrics and Gynecology*. Vol. 104 (5 Pt 1), Pp. 1051-1057.
- Kenyon, S., Boulvain, M., Neilson, J.P. (2003). Antibiotics for preterm premature rupture of membranes. *Cochrane Database of Systematic Reviews*. Issue 2. Art. No.: CD001058. DOI: 10.1002/14651858.
- Kesmodel, U., Wisborg, K., Olsen, S.F., Henriksen, T.B., and Secher, N.J. (2002). Moderate alcohol intake during pregnancy and the risk of still birth and death in the first year of life. *American Journal of Epidemiology*. Vol. 155, Issue 4, Pp 305-312.
- Khattree, R. and Naik, D.N. (2000). *Multivariate Data Reduction and Discrimination with SAS Software*. Cary, NC: SAS Institute Inc.
- Khong, T.Y., De Wolf, F., Robertson, W.B., and Brosens, I. (1986). Inadequate maternal vascular response to placentation in pregnancies complicated by pre-eclampsia and by small-for-gestational age infants. *British Journal of Obstetrics and Gynaecology*. Vol. 93, Pp. 1049-1059.
- Kirchengast, S., Hartmann, B., Schweppe, K.W., and Husslein, P. (1998). Impact of maternal body build characteristics on newborn size in two different European populations. *Human Biology; an International Record of Research*. Vol. 70, Issue 4, Pp. 761-774.
- Kish, L. (1965). *Survey Sampling*. New York: John Wiley & Sons.
- Klebanoff, M.A., Mednick, B.R., Schulsinger, C., Secher, N.J., and Shiono, P.H. (1998) Father's effect on infant birth weight. *American Journal of Obstetrics and Gynecology*. Vol. 178, Issue 5, Pp. 1022-1026.
- Kliegman, R.M., Rottman, C.J., and Behrman, R.E. (1990). Strategies for the prevention of low birth weight. *American Journal of Obstetrics and Gynecology*, Vol. 162, Issue 4, Pp. 1073-1083.
- Kliegman, R.M., and Das, U.G. (2002). Intrauterine growth retardation. In: Fanaroff AA, Martin RJ,

- editors. *Neonatal-Perinatal Medicine: Diseases of the Foetus and Infant*. 7<sup>th</sup> ed. New York: Mosby; Pp. 228-262.
- Kramer, M.S., Seguin, L., Lydon, J., et al. (2000). Socio-economic disparities in pregnancy outcome: why do the poor fare so poorly? *Paediatric and Perinatal Epidemiology*. Vol. 14, Issue 3, Pp. 194-210.
- Krieger, N., Chen, J., Waterman, P., et al. (2003). Choosing area based socioeconomic measures to monitor social inequalities in low birth weight and child lead poisoning: The Public Health Disparities Geocoding Project (US). *Journal of Epidemiology and Community Health*. Vol. 57, Pp. 186-199.
- Krige, D.G. (1951). *A Statistical Approach to Some Mine Valuations and Allied Problems at the Witwatersrand*. Master's thesis of the University of Witwatersrand.
- Krull, J.L. and MacKinnon, D.P. (1999). Multilevel mediation modeling in group-based intervention studies. *Evaluation Review*. Vol. 23 No. 4, Pp. 418-444.
- Kuh, D. and Ben-Shlomo, Y. (Eds.). (2004). *A Life Course Approach to Chronic Disease Epidemiology*. Oxford: Oxford University Press.
- Kuhnert, B.R., Kuhnert, P.M., Debanne, S., and Williams, T.G. (1987). The relationship between cadmium, zinc, and birth weight in pregnant women who smoke. *American Journal of Obstetrics and Gynecology*. Vol. 157, Issue 5, Pp. 1247-1251.
- Kulldorff, M. and Nagarwalla, N. (1995). Spatial disease clusters: detection and inference. *Statistics in Medicine*. Vol. 14, Pp. 799-810.
- Kulldorff, M. (1997). A spatial scan statistic. *Communications in Statistics: Theory and Methods*. Vol. 26, Pp. 1481-1496.
- Lane, S.D., Keefe, R.H., Rubinstein, R., Levandowski, B.A. et al. (2008). Structural violence, urban retail food markets, and low birth weight. *Health & Place*. Vol. 14, Issue 3, Pp. 415-423.
- Lang, J.M., Lieberman, E., and Cohen, A. (1996). A comparison of risk factors for preterm labor and term small-for-gestational-age birth. *Epidemiology*. Vol. 7, Pp. 369-376.
- Laudignon, N., Farri, E., Beharry, K., and Aranda, J.V. (1991). Rapid effects of hypoxia on the cerebrospinal fluid levels of adenosine and related metabolites in newborn and one-month-old piglets. *Biology of the Neonate*. Vol. 59, Issue 1, Pp. 54-59.
- Law, J. and Haining, R.P. (2004). A Bayesian approach to modeling binary data: the case of high-intensity crime areas. *Geographical Analysis*. Vol. 36, Issue 3, Pp. 197-216.
- Lawley, D.N. (1940). The estimation of factor loadings by the method of maximum likelihood. *Proceedings of Royal Society Edinburgh (A)*, Vol. 60, Pp. 64-82.
- Lawlor, D.A., Leon, D.A., and Davey, S.G. (2005). The association of ambient outdoor temperature

- throughout pregnancy and offspring birthweight: findings from the Aberdeen Children of the 1950s cohort. *BJOG: an International Journal of Obstetrics and Gynaecology*. Vol. 112, Issue 5, PP. 647-657.
- Lawson, A., Biggeri, A., Bohning, D., Lesaffre, E., Viel, J-F. and Bertollini, R. (eds.) (1999). *Disease Mapping and Risk Assessment for Public Health*. Chichester: Wiley.
- Leung, T.N., Chung, T.K., Madsen, G., McLean, M., Chang, A.M., and Smith, R. (1999). Elevated mid-trimester maternal corticotrophin-releasing hormone levels in pregnancies that delivered before 34 weeks. *British Journal of Obstetrics and Gynaecology*. Vol. 106, Issue 10, Pp. 1041-1046.
- Leviton, A., and Cowan, L. (2002). A review of the literature relating caffeine consumption by women to their risk of reproductive hazards. *Food and Chemical Toxicology*. Vol. 40, Issue 9, Pp. 1271-1310.
- Linn, R.L. (1968). A Monte Carlo approach to the number of factors problem. *Psychometrika*. Vol. 33, Pp. 37-71.
- Liszka, T. (1984). An interpolation method for an irregular net of nodes. *International Journal for Numerical Methods in Engineering*. Vol. 20, Issue 9, Pp. 1599-1612.
- Lui, K.J., Cumberland, W.G., and Kuo, L. (1996). An interval estimate for the intraclass correlation in beta-binomial sampling. *Biometrics*, Vol. 52, Issue 2, Pp. 412-425.
- Lukaszuk, S. (2004). A new concept of probability metric and its applications in approximation of scattered data sets. *Computational Mechanics*. Vol. 33, Pp. 299-304.
- Luke, B. (1994). Maternal-fetal nutrition. *Clinical Obstetrics and Gynecology*. Vol. 37, Issue 1, Pp.93-109.
- Luke, D.A. (2004). *Multilevel Modeling*. Thousand Oaks: SAGE Publications.
- Luo, Z.C., Kierans, W.J., Wilkins R., et al. (2004). Disparities in birth outcomes by neighbourhood income temporal trends in rural and urban areas, British Columbia. *Epidemiology*. Vol. 15, Issue 6, Pp. 679-686.
- Luo, Z.C., Wilkins, R., Kramer, M.S., et al. (2006). Effect of neighbourhood income and maternal education on birth outcomes: a population-based study. *Canadian Medical Association Journal*. Vol. 174, Issue 10, Pp. 1415-1420.
- Lumley, J., and Donohue, L. (2006). Aiming to increase birth weight: a randomised trial of pre-pregnancy information, advice and counselling in inner-urban Melbourne. *BMC Public Health*. Vol. 6, Pp. 299.
- Lundsberg, L.S., Bracken, M.B., and Saftlas, A.F. (1997). Low-to-moderate gestational alcohol use and intrauterine growth retardation, low birthweight, and preterm delivery. *Annals of*

- Epidemiology*. Vol. 7, Issue 7, Pp. 498-508.
- Lydon, J., Dunkel-Schetter C., Cohan, C.L., Perce, T. (1996). Pregnancy decision making as a significant life event: a commitment approach. *Journal of Personality and Social Psychology*. Vol. 71, Pp. 141-151.
- Lynch, J.W., Kaplan, G.A., and Salonen, J.T. (1997). Why do poor people behave poorly? Variation in adult health behaviours and psycho-social characteristics by stages of the socioeconomic lifecourse. *Social Science and Medicine*. Vol. 44, Issue 6, Pp. 809-819.
- Lynch, J.W., Davey Smith, G., Kaplan, G. A. and House, J.S. (2000). Income inequality and mortality: Importance to health of individual income, psychosocial environment, or material conditions. *British Medical Journal*. Vol. 320, Pp. 1200-1224.
- MacCredie, M., Williams, S., and Coates, M. (1999). Cancer mortality in East and Southeast Asian migrants to New South Wales, Australia, 1975-1995. *British Journal of Cancer*. Vol. 79, Issue 7-8, Pp. 1277-1282.
- Macintyre, S. and Ellaway, A. (1998). Social and local variations in the use of urban neighbourhoods: a case study in Glasgow. *Health and Place*. Vol. 4, Issue 1, Pp. 91-94.
- Mackenbach, J.P., Looman, C.W., and Kunst, A.E. (1993), Air pollution, lagged effects of temperature, and mortality: The Netherlands, 1979-87. *Journal of Epidemiological Community Health*. Vol. 47, Pp. 121-126.
- MacKinnon, D. P., and Dwyer, J.H. (1993). Estimating mediated effects in prevention studies. *Evaluation Review*. Vol. 17, Issue 2, Pp. 144-158.
- MacKinnon, D. P., Fairchild, A. J., and Fritz, M. S. (2007). Mediation analysis. *Annual Review of Psychology*. Vol. 58, Pp. 593 -614.
- MacMillan, H.L., and Wathen, C.N. (2003). Violence against women: integrating the evidence into clinical practice. *Canadian Medical Association Journal*. Vol. 169, Issue 6, Pp. 570-571.
- Magnus, P., Bakketeig, L.S., and Skjaerven, R. (1993). Correlations of birth weight and gestational age across generations. *Annals of Human Biology*. Vol. 20, Issue 3, Pp. 231-238.
- Malmstrom, M., Sundquist, J., Bajekal, M., and Johansson, S.E. (1999). Ten-years trends in all-cause mortality and coronary heart disease mortality in socio-economically diverse neighbourhoods. *Public Health*. Vol. 113, Pp. 279-284.
- Mantel, N. (1967). The detection of disease clustering and a generalised regression approach. *Cancer Research*. Vol. 27, Pp. 209-220.
- Marshall, R. (1991). A review of methods for the statistical analysis of spatial patterns of disease. *Journal of the Royal Statistical Society, Series A*. Vol. 154, 421-441.
- Mariscal, M., Palma, S., Llorca, J., Perez-Iglesias, R., Pardo-Crespo, R., and Delgado- Rodriguez, M.

- (2006). Pattern of alcohol consumption during pregnancy and risk for low birth weight. *Annals of Epidemiology*. Vol. 16, Issue 6, Pp. 432-438.
- Marmot, M., and Wilkinson, R.G. (2006). *Social Determinants of Health*. 2<sup>nd</sup> edition. New York: Oxford University Press.
- Martens, P. and Rotmans, J. (eds.). (2002) *Transitions in a Globalising World*. Lisse: Swets &Zeitlinger.
- Martin, J. A. et al. (2007). Births: Final Data for 2005, *National Vital Statistics Reports* 56, 6. Hyattsville, Maryland: National Center for Health Statistics.
- Musterd, S. and Andersson, R. (2005). Housing Mix, Social Mix, and Social Opportunities. *Urban Affairs Review*. vol. 40, Issue 6, Pp. 761-790.
- May, J. (1958). *The Ecology of Human Disease*. New York: MD.
- McCormick, M.C. (1985). The contribution of low birth weight to infant mortality and childhood morbidity. *The New England Journal of Medicine*. Vol. 312, Issue 2, Pp. 82-90.
- McDonald, S.D., Murphy, K., Beyene, J., and Ohlsson, A. (2005). Perinatal outcomes of singleton pregnancies achieved by in vitro fertilization: a systematic review and metaanalysis. *Journal of Obstetrics and Gynaecology Canada*. Vol. 27, Issue 5, Pp. 449-459.
- McGuffin, P., Riley, B., and Plomin, R. (2001). Toward behavioural genomics. *Science*. Vol. 291, Pp. 1232-1233.
- Meade, M. and Earickson, R. (2000). *Medical Geography*, 2<sup>nd</sup> edition. New York: Guilford.
- Meng, G. (2005). Spatial Segregation of Housing Quality in Lima, Peru. Master's thesis of the University of Waterloo.
- Metalsky, G.I., Joiner T.E., Hardin, T.S., and Abramson L.Y. (1993). Depressive reactions to failure in a naturalistic setting: a test of the hopelessness and self-esteem theories of depression. *Journal of Abnormal Psychology*. Vol. 102, Issue 1, Pp. 101-109.
- Millar, L.K., and Cox, S.M. (1997). Urinary tract infections complicating pregnancy. *Infectious Disease Clinics of North America*. Vol. 11, Issue 1, Pp. 13-26.
- Miller, J.L. (1978). *Living systems*. New York: McGraw Hill.
- Misra, D.P., and Nguyen, R.H. (1999). Environmental tobacco smoke and low birth weight: a hazard in the workplace? *Environmental Health Perspectives*. Vol. 107 (Suppl 6), Pp. 897-904.
- Mitchell, R., Dorling, D. and Shaw, M. (2000a). *Inequalities in Life and Death: What if Britain Were More Equal?* Bristol: The policy Press.
- Mitchell, R., Gleave, S. Bartley, M. and Wiggins, R. (2000b). Do attitude and area influence health? A multilevel approach to health inequalities. *Health and Place*. Vol. 6, Issue 2, Pp. 67-69.
- Molin, A. (1993). Risk of damage to the cervix by dilatation for first-trimester-induced abortion by suction aspiration. *Gynecologic and Obstetric Investigation*. Vol. 35, Issue 3, Pp. 152-154.

- Moore, M.L., Meis, P.J., Ernest, J.M., Wells, H.B., Zaccaro, D.J., and Terrell, T. (1998). A randomized trial of nurse intervention to reduce preterm and low birth weight births. *Obstetrics and Gynecology*. Vol. 91(5 Pt 1), Pp. 656-61.
- Moran, P. A. P. (1948). The interpretation of statistical maps. *Journal of Royal Statistical Society, Series B*. Vol. 10, Pp. 243-251.
- Moran, P. A. P. (1950). Notes on continuous stochastic phenomena. *Biometrika*. Vol. 37, Pp. 17-23.
- Morency, A.M., and Bujold, E. (2007). The effect of second-trimester antibiotic therapy on the rate of preterm birth. *Journal of Obstetrics and Gynaecology Canada: JOGC*. Vol. 29, Issue 1, Pp. 35-44.
- Morenoff, J.D. (2003). Neighborhood mechanisms and the spatial dynamics of birth weight. *American Journal of Sociology*. Vol. 108, Issue 5, Pp. 976-1017.
- Morone, J.A. and Jacobs, L.R. (2005). *Healthy, Wealthy, and Fair: Health Care and the Good Society*. Oxford: Oxford University Press.
- Murphy, C.C., Schei, B., Myhr, T.L., and Du Mont, J. (2001). Abuse: a risk factor for low birth weight? A systematic review and meta-analysis. *Canadian Medical Association Journal*. Vol. 164, Issue 11, Pp. 1567-1572.
- Murphy, V.E., Clifton, V.L., and Gibson, P.G. (2006). Asthma exacerbations during pregnancy: incidence and association with adverse pregnancy outcomes. *Thorax*. Vol. 61, Issue 2, Pp. 169-176.
- Naeye, R.L., and Peters, E.C. (1982). Working during pregnancy: effects on the fetus. *Pediatrics*. Vol. 69, Issue 6, Pp. 724-727.
- Naimi, T.S., Lipscomb, L.E., Brewer, R.D., and Gilbert, B.C. (2003). Binge drinking in the preconception period and the risk of unintended pregnancy: implications for women and their children. *Pediatrics*. Vol. 111 (5 Pt 2), Pp. 1136-1141.
- Needleman, C. (1997). Applied epidemiology and environmental health: emerging controversies. *American Journal of Infection Control*. Vol. 25, Issue 3, Pp. 262-274.
- Neerhof, M.G. (1995). Causes of intrauterine growth restriction. *Clinics in Perinatology*. Vol. 22, Issue 2, Pp. 375-85.
- Nelder, J. A. and Pregibon, D. (1987). An extended quasilielihood function. *Biometrika*. Vol. 74, Pp. 221-232.
- Neuhaus, J.O. and Wrigley, C. (1954). The quartimax method: an analytical approach to orthogonal simple structure. *British Journal of Statistical Psychology*. Vol. 7, Pp. 81-91.
- Newburn-Cook, C.V., and Onyskiw, J.E. (1995). Is older maternal age a risk factor for preterm birth and foetal growth restriction? A systematic review. *Health Care for Women International*. Vol.

- 26, Issue 9, Pp. 852-875.
- Nurminen, T. (1995). Female noise exposure, shift work, and reproduction. *Journal of Occupational and Environmental Medicine*. Vol. 37, Issue 8, Pp. 945-950.
- O'Campo, P., Xue, X., Wang, M.C., and Caughy, M. (1997). Neighborhood risk factors for low birthweight in Baltimore: a multilevel analysis. *American Journal of Public Health*. Vol. 87, Issue 7, Pp. 1113-1118.
- Oden, N. (1995). Adjusting Moran's I for population density. *Statistics in Medicine*. Vol. 14, Pp. 17-26.
- Olsen, S.F., Secher, N.J., Tabor, A., Weber, T., Walker, J.J., and Gluud, C. (2000). Randomised clinical trials of fish oil supplementation in high risk pregnancies. Fish Oil Trials In Pregnancy (FOTIP) Team. *BJOG: an International Journal of Obstetrics and Gynaecology*. Vol. 107, Issue 3, Pp. 382-395.
- Osler, M., Gerdes, L.U., Davidsen, M., Bronnum-Hansen, H., Madsen, M., Jorgensen, T., and Schroll, M. (2000). Socioeconomic status and trends in risk factors for cardiovascular diseases in the Danish MONICA population, 1982-1992. *Journal of Epidemiology and Community Health*. Vol. 54, Pp. 108-113.
- Openshaw, S., and Taylor, P. (1979). A million or so correlation coefficients: Three experiments on the modifiable areal unit problem. In N. Wrigley (Ed.), *Statistical Applications in the Spatial Sciences*. London: Pion. Pp. 127-144.
- Osypuk, T.L., and Acevedo-Garcia, D. (2008). Are racial disparities in preterm birth larger in hypersegregated areas? *American Journal of Epidemiology*. Vol. 167, No. 11, Pp. 1295-1304.
- Overall, A. and Nichols, R. (2001). A method for distinguishing consanguinity and population substructure using multilocus genetic data. *Molecular Biology and Evolution*. Vol. 18, Issue 11, Pp. 2048-2056.
- Pacione, M. (Ed.) (1987). *Social Geography: Progress and Prospect*. London: CROOM HELM.
- Pak, L.L., Reece, E.A., and Chan, L. (1998). Is adverse pregnancy outcome predictable after blunt abdominal trauma? *American Journal of Obstetrics and Gynecology*. Vol. 179, Issue 5, Pp. 1140-1144.
- Pallitto, C.C., Campbell, J.C., and O'Campo, P. (2005). Is intimate partner violence associated with unintended pregnancy? A review of the literature. *Trauma, Violence & Abuse*. Vol. 6, Issue 3. Pp. 217-235.
- Paneth, N.S. (1995). The problem of low birth weight. *The Future of Children*. Vol. 5, Pp 19-34.
- Papiernik, E., and Goffinet, F. (2004). Prevention of preterm births, the French experience. *Clinical Obstetrics and Gynecology*. Vol. 47, Issue 4, Pp. 755-767.

- Parkes, M., Panelli, R. and Weinstein, P. (2003). Converging paradigms for environmental health theory and practice. *Environmental Health Perspectives*. Vol. 11, No. 5, Pp. 669-675.
- Pastrakuljic, A., Derewlany, L.O., and Koren, G. (1999). Maternal cocaine use and cigarette smoking in pregnancy in relation to amino acid transport and foetal growth. *Placenta*. Vol. 20, Issue 7, Pp. 499-512.
- Pattenden, S., Dolk, H., and Vrijheid, M. (1999). Inequalities in low birth weight: Parental social class, area deprivation, and “lone mother” status. *Journal of Epidemiology and Community Health*. Vol. 53, Issue 6, Pp. 355-358.
- Penchansky, R. and Thomas, J. (1981). The concept of access: definition and relationship to consumer satisfaction. *Medical Care*. Vol. XIX, Issue 2, Pp. 127-140.
- Peter, R., and Siegrist, J. (1997). Chronic work stress, sickness absence, and hypertension in middle managers: general or specific sociological explanations? *Social Science and Medicine*. Vol. 45, Issue 7, Pp. 1111-1120.
- Pfeffermann, D. (2002). Small area estimation: New developments and directions. *International Statistical Review*, 2002. Vol. 70, Pp. 125-143.
- Pickett, K.E., Collins, J.W., Jr., Masi, C.M., and Wilkinson, R.G. (2005). The effects of racial density and income incongruity on pregnancy outcomes. *Social Science & Medicine*. Vol. 60, Issue 10, Pp. 2229-2238.
- Plessinger, M.A. (1998). Prenatal exposure to amphetamines. Risks and adverse outcomes in pregnancy. *Obstetrics and Gynecology Clinics of North America*. Vol. 25, Issue 1, Pp. 119-138.
- Pollack, R.N., and Divon, M.Y. (1992). Intrauterine growth retardation: definition, classification, and etiology. *Clinical Obstetrics and Gynecology*. Vol. 35, Issue 1, Pp. 99-107.
- Prasad, M., and Al-Taher, H. (2002). Maternal height and labour outcome. *Journal of Obstetrics and Gynaecology*. Vol. 22, Issue 5, Pp. 513-515.
- Public Health Agency of Canada (2008). Canadian Perinatal Health Report, 2008 Edition. Ottawa.
- Putnam, R. (1993). The prosperous community: Social capital and community life. *American Prospect*. Vol. 35, Pp. 42.
- Raatikainen, K., Heiskanen, N., and Heinonen, S. (2005). Marriage still protects pregnancy. *BJOG: an International Journal of Obstetrics and Gynaecology*. Vol. 112, Issue 10, Pp. 1411-1416.
- Ramin, S.M., Vidaeff, A.C., Yeomans, E.R., and Gilstrap, L.C. (2006). Chronic renal disease in pregnancy. *Obstetrics and Gynecology*. Vol.108, Issue 6, Pp. 1531-1539.
- Ramsay, J., Richardson, J., Carter, Y.H., Davidson, L.L., and Feder, G. (2002). Should health professionals screen women for domestic violence? Systematic review. *British Medical Journal*. Vol.325, Issue 7359, Pp. 314.



- Ramsay, P.S., and Goldenberg, R.L. (2002). Obstetric management of prematurity. In: Fanaroff AA, Martin RJ, editors. *Neonatal-Perinatal Medicine: Diseases of the foetus and infant*. 7<sup>th</sup> ed. New York: Mosby. Pp. 287-319.
- Rao, C.R. (1955). Estimation and tests of significance in factor analysis. *Psychometrika*. Vol. 20, Pp. 93-111.
- Rao, C.R. (1964). The use of interpretation of principal components in applied research. *Sankhya, A*, Vol. 26, Pp. 329-358.
- Raphael, D. (ed.) (2004). *Social Determinants of Health: Canadian Perspectives*. Toronto: Canadian Scholars' Press.
- Reagan, P.B., and Salsberry, P.J. (2005). Race and ethnic differences in determinants of preterm birth in the USA: broadening the social context. *Social Science & Medicine*. Vol. 60, Issue 10, Pp. 2217-2228.
- Ridout, M.S., Demetrio, C.G.B., and Firth, D. (1999). Estimating intraclass correlation for binary data. *Biometrics*. Vol. 55, Issue 1, Pp. 137-148.
- Ritz, B., Yu, F., Chapa, G., and Fruin, S. (2000). Effect of air pollution on preterm birth among children born in Southern California between 1989 and 1993. *Epidemiology*. Vol. 11, Pp. 502-511.
- Ritz, B., Yu, F., Fruin, S., Chapa, G., Shaw, G.M., and Harris, J.A. (2002). Ambient air pollution and risk of birth defects in southern California. *American Journal of Epidemiology*. Vol. 155, Pp. 17-25.
- Roberts, G. (1996). Markov chain concepts related to sampling algorithms. In W. Gilks, S. Richardson and D. Spiegelhalter, Editors, *Markov Chain Monte Carlo in Practice*. Chapman & Hall, London, Pp. 45-57.
- Romero, R., Gomez, R., Chaiworapongsa, T., Conoscenti, G., Kim, J.C., and Kim, Y.M. (2001). The role of infection in preterm labour and delivery. *Paediatric and Perinatal Epidemiology*. Vol. 15 (Suppl 2), Pp. 41-56.
- Roth, J., Hendrickson, J., Schilling, M., and Stowell, D.W. (1998). The risk of teen mothers having low birth weight babies: implications of recent medical research for school health personnel. *The Journal of School Health*. Vol. 68, Issue 7, Pp. 271-275.
- Rushton, L. (2003). Health hazards and waste management. *British Medical Bulletin*. Vol. 68, Pp. 183-197.
- Rutledge, J. (2000). Perimplantation teratology and the placenta. *Teratology*. Vol. 61, Pp. 246-247.
- Sack, R. (1986). *Human Territoriality, its Theory and History*. Cambridge: Cambridge University Press.

- Sammons, M.N. (1981). Battered and pregnant. *American Journal of Maternal/Child Nursing*. Vol. 6, Pp. 246-50.
- Sampson, R.J. (1997). Neighborhoods and violent crime: A multilevel study of collective efficacy. *Science*. Vol. 227, Pp. 918-924.
- Sanger, M. and Sinclair, S. (Eds.). (2004). *Putting Health First: Canadian health care reform in a globalizing world*. Ottawa: Canadian Centre for Policy Alternatives.
- Santos, J.F., Ribeiro, R.M., Rossi, P., Haddad, J.M., Guidi, H.G., Pacetta, A.M., et al. (2002). Urinary tract infections in pregnant women. *International Urogynecology Journal and Pelvic Floor Dysfunction*. Vol. 13, Issue 3, Pp. 204-9.
- Sastry, N., and Hussey, J.M. (2003). An investigation of racial and ethnic disparities in birth weight in Chicago neighborhoods. *Demography*. Vol. 40, Issue 4, Pp. 701-725.
- Scholl, T.O., Hediger, M.L., and Schall, J.I. (1996). Mechanisms for maternal age differences in birth weight. *American Journal of Epidemiology*. Vol. 143, Issue 8, Pp. 842-843.
- Scott, H.M., O'Connor, K.O., and Carr, P.A. (2001). Evaluation of a community-wide preterm birth prevention program. *Canadian Journal of Public Health*. Vol. 92, Issue 5, Pp. 328-330.
- Seguin, L., Potvin, L., Loiselle, and J., St-Denis, M. (1995). Chronic stressors, social support, and depression during pregnancy. *Obstetrics and Gynecology*. Vol. 85, Pp. 583-589.
- Shaw, M. Dorling, D. and Mitchell, R. (2002). *Health, Place and Society*. Great Britain: Pearson Education Limited.
- Shepard, D. (1968). A two-dimensional interpolation function for irregularly-spaced data. *Proceedings of the 1968 ACM National Conference*. Pp. 517-524.
- Siegrist, J. (2000). Place, social exchange and health: proposed sociological framework. *Social Science and Medicine*. Vol. 51, Pp. 1283-1293.
- Simpson, J.L. (1993). Are physical activity and employment related to preterm birth and low birth weight? *American Journal of Obstetrics and Gynecology*. Vol. 168, Issue 4, Pp. 1231- 1238.
- Smets, P. and Salman, T. (2008). Countering Urban Segregation: Theoretical and Policy Innovations from around the Globe. *Urban Studies*. Vol. 45, Pp. 1307-1332.
- Smith, J. L. (2006) Mixed-income communities: designing out poverty or pushing out the poor? in: L. Bennett, J. L. Smith and P. A. Wright (Eds) *Where are Poor People to Live? Transforming Public Housing Communities*, Chapter 10. Armonk, NY: M. E. Sharpe.
- Sobel, M. E. (1982). Asymptotic confidence intervals for indirect effects in structural equation models. In S. Leinhardt (Ed.), *Sociological Methodology*. Washington DC: American Sociological Association.
- Sosa, C., Althabe, F., Belizan, J., and Bergel, E. (2004). Bed rest in singleton pregnancies for

- preventing preterm birth. *Cochrane Database of Systematic Reviews*. Issue 1. Art. No.: CD003581. DOI: 10.1002/14651858.
- Spiegelhalter, D.J., Best, N.G., Carlin, B.P. and van der Linde, A. (2002) Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B*. Vol. 64, Pp. 583-640.
- Sprague, A.E. (2004). The evolution of bed rest as a clinical intervention. *Journal of Obstetric, Gynecologic, and Neonatal Nursing*. Vol. 33, Issue 5, Pp. 542-549.
- Stan, C., Boulvain, M., Hirsbrunner-Amagbaly, P., and Pfister, R. (2002). Hydration for treatment of preterm labour. *Cochrane Database of Systematic Reviews*. Issue 2. Art. No.: CD003096. DOI: 10.1002/14651858.
- Statistics Canada (2006). Ethnocultural Portrait of Canada - Data table, Census 2006.
- Steptoe, A. (2000). Stress, social support and cardiovascular activity over the working day. *International Journal of Psychophysiology*. Vol. 37, Issue 3, Pp. 299-308.
- Stern, H.S. and Cressie, N.A. (1999). Inference for extremes in disease mapping. In *Disease Mapping and Risk Assessment for Public Health*. A. Lawson, A. Biggeri, D. Bohning, E. Lesaffre, J-F. Viel and R. Bertollini (eds.), Chichester: Wiley, Pp. 63-84.
- Stein, M.L. (1999). *Statistical Interpolation of Spatial Data: Some Theory for Kriging*. Springer, New York.
- Stone, C. (1993). Urban regimes and the capacity to govern: a political economy approach. *Journal of Urban Affairs*. Vol. 15, Issue 1, Pp. 1-28.
- Sturchler, D., Menegoz, F., and Daling, J. (1986). Reproductive history and intrapartum fever. *Gynecologic and Obstetric Investigation*. Vol. 21, Issue 4, Pp.182-186.
- Subramanian, S.V., Jones, K., and Duncan, C. (2003). Multilevel methods for public health research. In Kawachi, I and Berkman, L.F. (Eds.). *Neighborhoods and Health*. Oxford: Oxford University Press.
- Susser, M. (1985). Epidemiology in the United States after World War II: the evolution of technique. *Epidemiologic Reviews*. Vol. 7, Pp. 147-177.
- Tango, T. (1995). A class of test for detecting ‘general’ and ‘focused’ clustering of rare diseases. *Statistics in Medicine*. Vol. 14, Pp. 2323-2334.
- Tarlatzis, B.C., and Zepiridis, L. (2003). Perimenopausal conception. *Annals of the New York Academy of Sciences*. Vol. 997, Pp. 93-104.
- Tekesin, I., Eberhart, L.H., Schaefer, V., Wallwiener, D., and Schmidt, S. (2005). Evaluation and validation of a new risk score (CLEOPATRA score) to predict the probability of premature delivery for patients with threatened preterm labour. *Ultrasound in Obstetrics & Gynaecology*.

- Vol. 26, Issue 7, Pp. 699-706.
- Ten Have, T.R., Ratcliffe, S.J., Reboussin, B.A., and Miller, M.E. (2004). Deviations from the population-averaged versus cluster specific relationship for clustered binary data. *Statistical Methods in Medical Research*. Vol. 13, Issue 3, Pp. 16.
- Theorell, T. (2000). Working conditions and health. In Berkman, L.F. and Kawachi, I. (eds.). *Social Epidemiology*. Oxford: Oxford University Press.
- Tierney, L. (1996). Introduction to general state-space Markov chain theory. In W. Gilks, S. Richardson and D. Spiegelhalter, Editors, *Markov Chain Monte Carlo in Practice*. Chapman & Hall, London, Pp. 45-57.
- Townsend, P., Davidson, N., and Whitehead, M. (1986). *The Black Report and the Health Divide*. Harmondsworth: Penguin.
- Townsend, P., Davidson, N., and Whitehead, M. (1992). *Inequalities in Health: The Black Report and the Health Divide*. New York: Penguin.
- Turnbull, B., Iwano, E., Burnett, W., Howe, H. and Clark, L. (1990). Monitoring for clusters of disease: application to leukaemia incidence in upstate New York. *American Journal of Epidemiology*. Vol. 132: supplement, Pp. 136-143.
- Uchida, N., Suzuki, A., Ohyama, K., Bessho, T., and Toyoda, H. (2006). Secretion of bioactive interleukin-6 and tumor necrosis factor-alpha proteins from primary cultured human foetal membrane chorion cells infected with influenza virus. *Placenta*. Vol. 27(6-7), Pp. 678-690.
- United Nations Children's Fund and World Health Organization, *Low Birthweight: Country, Regional and Global Estimates* (New York, New York: UNICEF, 2004).
- Urquia, M.L., Frank, J.W., Glazier, R.H., and Moineddin, R. (2007). Birth outcomes by neighbourhood income and recent immigration in Toronto. *Statistics Canada Health Reports*. Vol. 18, No. 4.
- Utsey, S.O., Chae, M.H., Brown, C.F., and Kelly, D. (2002). Effect of ethnic group membership on ethnic identity, race-related stress, and quality of life. *Cultural Diversity & Ethnic Minority Psychology*. Vol. 8, Issue 4, Pp. 366-377.
- Valero De Bernabé J. et al. (2004). Risk Factors for Low Birth Weight: A Review, *European Journal of Obstetrics, Gynecology, and Reproductive Biology*. Vol. 116, Issue 1, Pp. 3-15.
- Velicer, W.F. (1977). An empirical comparison of the similarity of principal component, image and factor patterns, *Multivariate Behavioral Research*. Vol. 12, Pp. 3-22.
- Villar, J., Carroli, G., Wojdyla, D., Abalos, E., Giordano, D., Ba'aqeel, H., et al. (2006). Preeclampsia, gestational hypertension and intrauterine growth restriction, related or independent conditions? *American Journal of Obstetrics and Gynecology*. Vol. 194, Issue 4, Pp.

921-31.

- Virji, S.K., and Talbott, E.O. (1990). The relationship between occupational classification and low birth weight in a national sample of white married mothers. *International Archives of Occupational and Environmental Health*. Vol. 62, Issue 5, Pp. 351-356.
- Wackernagel, H. (2003). *Multivariate Geostatistics*. Springer.
- Wadhwa, P.D., Culhane, J.F., Rauh, V., et al. (2001). Stress, infection and preterm birth: a biobehavioural perspective. *Paediatric and Perinatal Epidemiology*. Vol. 15 (Suppl), Pp. 17-29.
- Waldron, I., Hughes, M.E., and Brooks, T.L. (1996). Marriage protection and marriage selection - prospective evidence for reciprocal effects of marital status and health. *Social Science & Medicine*. Vol. 43, Issue 1, Pp. 113-123.
- Wallace, R. and Wallace, D. (1997). Socio-economic determinants of health: community marginalisation and the diffusion of disease and disorder in the United States. *British Medical Journal*. Vol. 314, Pp. 1341-1345.
- Walpole, I., Zubrick, S., and Pontre J. (1990). Is there a foetal effect with low to moderate alcohol use before or during pregnancy? *Journal of Epidemiology and Community Health*. Vol. 44, Pp 297-301.
- Walsh, R.A., Lowe, J.B., and Hopkins, P.J. (2001). Quitting smoking during pregnancy. *Medical Journal of Australia*. Vol. 175, Issue 6, Pp. 320-323.
- Wang, X., Ding, H., Ryan, L., and Xu, X. (1997). Association between air pollution and low birth weight: a community-based study. *Environmental Health Perspectives*. Vol. 105, Pp. 514-520.
- Wathen, C.N., and MacMillan, H.L. (2003). Interventions for violence against women: scientific review. *The Journal of the American Medical Association*. Vol.289, Issue 5, Pp. 589-600.
- Watkins, M.L., Rasmussen, S.A., Honein, M.A., Botto, L.D., and Moore, C.A. (2003). Maternal obesity and risk for birth defects. *Pediatrics*. Vol. 111 (5 Part 2), Pp. 1152-1158.
- Wellman, B. and Leighton, B. (1979). Networks, neighbourhoods and communities. *Urban Affairs Quart*. Vol. 14, Pp. 363-390.
- White, D.E., Fraser-Lee, N.J., Tough, S., and Newburn-Cook, C.V. (2006). The content of prenatal care and its relationship to preterm birth in Alberta, Canada. *Health Care for Women International*. Vol. 27, Issue 9, Pp. 777-792.
- Wilson, W.J. (1996). *When Work Disappears*. New York: Knopf.
- Wilkins, R., Sherman, G.J., and Best, P.A. (1991). Birth outcomes and infant mortality by income in urban Canada, 1986. *Health Reports*. Vol. 3, Issue 1, Pp. 7-31.
- Wilkinson, R.G. and Pickett, K.E. (2006). Income inequality and population health: A review and explanation of the evidence. *Social Science & Medicine*. Vol. 62, Issue 7, Pp. 1768-1784.

- Windham, G.C., Eaton, A., and Hopkins, B. (1999). Evidence for an association between environmental tobacco smoke exposure and birthweight: a meta-analysis and new data. *Paediatric and Perinatal Epidemiology*. Vol. 13, Issue 1, Pp. 35-57.
- Wolfinger, R. and O'Connell, M. (1993), Generalized linear mixed models: a pseudo-likelihood approach. *Journal of Statistical Computation and Simulation*. Vol. 4, Pp. 233-243.
- Wonnacott, T.H. and Wonnacott, R.J. (1981). *Regression, a Second Course in Statistics*. New York: Wiley.
- World Health Organization (1998). *Health Promotion Glossary*, WHO.
- Xu, X., Ding, H., and Wang, X. (1995). Acute effects of total suspended particulate and sulfur dioxides on preterm delivery: a community-based cohort study. *Archives of Environmental Health*. Vol. 50, Pp. 407-415.
- Yamamoto, E. and Yanagimoto, T. (1997). Moment estimators for the binomial distribution. *Journal of Applied Statistics*. Vol. 19, Pp. 273-283.
- Yartey, J.E. (2006). Malaria in pregnancy: access to effective interventions in Africa. *International Journal of Gynaecology and Obstetrics*. Vol. 94, Issue 3, Pp. 364-373.
- Young, T.K. (2005). *Population Health: concepts and methods*. Oxford: Oxford University Press.
- Yost, N.P., Bloom, S.L., McIntire, D.D., and Leveno, K.J. (2005). Hospitalization for women with arrested preterm labor: a randomized trial. *Obstetrics and Gynecology*. Vol. 106, Issue 1, Pp. 14-18.
- Zaren, B., Cnattingius, S., and Lindmark, G. (1997). Foetal growth impairment from smoking -- is it influenced by maternal anthropometry? *Acta Obstetrica et Gynecologica Scandinavica Supplementum*. Vol. 165, Pp. 30-34.
- Zeger, S.L., Liang, K.Y., and Albert, P.S. (1988). Models for longitudinal data: a generalized estimating equation approach. *Biometrics*. Vol. 44, Pp. 1049-1060.
- Zou, G. and Donner, A. (2004). Confidence interval estimation of the intraclass correlation coefficient for binary outcome data. *Biometrics*. Vol. 60, Pp. 807-811.
- Zuckerman, B.S., and Hingson, R. (1986). Alcohol consumption during pregnancy: a critical review. *Developmental Medicine and Child Neurology*. Vol. 28, Pp. 649-661.